

A new ball-park for research?

Nature will not be taking sides in the forthcoming British general election on 11 June, which is appropriate but which may be thought to require explanation.

NOBODY is particularly surprised that the British government has called a general election roughly a year before it would be constitutionally compelled to do so. The second Thatcher government was never as confident as the first, most probably for the simple reason that new brooms sweep more energetically than those that are slightly worn. For much of its second term, the government has been battling with nothing more dramatic than the solution of problems left over from its first spell in office — the restoration of sound money, the reduction of unemployment and the efficient management of ever-day affairs. This recipe for dullness has not been uniformly successful. From time to time, there have been rows about the sale of motor-car and helicopter manufacturers to foreign investors and about the conduct of retired agents of the British intelligence services. But the second Thatcher administration seems to have made a genuine attempt to make the role of government less obtrusive. In recent months, its fortitude appears to have been rewarded by the improvement of economic indicators and a downturn in the previously rising trend on unemployment. Who can blame it for making a dash for a new mandate at the polling-booths?

But is this not the very administration that has superintended the dramatic, some say terminal, decline of the vigour of the British research enterprise, and which on that account alone might be denied a third term? That argument is over-simple. For one thing, the decline predates 1979. For another, and to the extent that one of the chief attributes of government policy towards civil research and the universities has been budgetary meanness, there is much in what might be its rejoinder to its critics — that the need to reduce public spending was not, at the outset, sufficiently well appreciated. Third, the British research community is itself partly to blame for what has happened; in the management of quite substantial resources, to which there have been few strings attached, the community has often been less imaginative and less decisive than it should have been. The chief complaint against the government is that it has been insensitive in its management of its research enterprise, feckless in its choice of managers thereof and indifferent to the danger that management by the repeated redefinition of objectives is a recipe for the erosion of morale. There are few heroes in the tale of what has happened in the past few years. Equally, there is no assurance that a government of a different stripe would have done markedly better.

Harassment

What has gone wrong with British science is mostly that attempts by successive governments to coerce the system of universities and research councils into the paths of economy and productivity have come to seem like harassment to those working in laboratories. Having decreed a 13.5 per cent cut in university budgets (at the end of 1980), the government found itself compelled to relent before the period of retrenchment was over (but not before many able people had been put out to grass), spending an extra £30 million on a 'new blood' scheme to fill the gaps. From time to time, the government has been breathing fire about the high cost of high-energy physics and the British membership of the European collaborative laboratory (CERN) at

Geneva (see p. 178), but has kept those concerned on tenterhooks by asking for successive inquiries into the problem. Sometimes there has been sheer bungling, as in this year's fracas over university salaries; somebody had forgotten that a 24 per cent pay increase for academics would cost the research councils the best part of £20 million.

Ministers

In the government's defence, it has been unlucky in its ministers. Its first Secretary of State for Education and Science, Mr Mark Carlisle, is remembered only for having announced the planned university cuts on the eve of the Christmas recess of the British parliament in 1980. His successor, Sir Keith Joseph, is vividly but not warmly remembered; his disastrous discussion paper on higher education last year showed the government to be not so much an imaginative architect of important social institutions as a kind of dull quantity surveyor anxious that not a single extra brick would be used. The present incumbent of the office, Mr Kenneth Baker, is more engaging and decisive; his white paper on higher education, nominally the flesh in which Sir Keith Joseph's skeleton is clothed, manages not to refer to the earlier document, which is just as well. But there will be ructions of a kind when British academics of all kinds have taken in the implications of the proposed new mechanism for supporting universities out of public funds (see below).

Like many governments elsewhere, the British government has been seeking a way of winning greater economic benefit from what it spends on research. The aim is laudable, but the method has not been thought through. It is doubtful whether forcing academics short of funds into the arms of industrial contract-writers will revive the British economy, although the government can boast of a healthy growth in the number of small companies spawned by academic research now that the City of London has learned what technology is about. Last month's decision that spending on defence research should be trimmed would have been more welcome if it had been clearer how the funds thus saved will be spent elsewhere. There is little doubt that this mishmash of decisions would have been more coherent if the government had followed the advice of the House of Lords in 1981, and had given science and technology an explicit place in its own political structure.

Those who are tempted to take this muddled record as a sufficient reason for voting against the government at the general election on 11 June should, however, give the matter careful thought. The government has indeed kept Mrs Thatcher's promise in 1980 that the budgets of the research councils would be "protected"; collectively, they have even increased a little in real terms. (The problems that have arisen stem from the secular decline of the value of sterling and the requirement that the councils should do extra things, from the support of information technology to the firing of staff members before retirement age.) The government has also succeeded in its declared aim of bringing inflation under control; if it had not, the British scientific enterprise would be in a greater mess.

Readers must judge for themselves whether the plans of the opposition parties are more convincing; both the Labour Party



and the Liberal-SDP Alliance claim that their hearts are better placed. An account of Labour policy will be found in *Nature* earlier this year (326, 437; 1987). Important though a government's policy on research may be, British voters will rightly say that the absorbing issue at this election is that of which of the two opposition parties comes out second to the government, predicted by the opinion polls to win. Journals such as this, concerned with news of the unfolding of research and thus with its efficient management, are in some difficulty. Broad questions such as the realignment of the opposition parties in just one small country would be baffling for the majority of readers, and are in any case irrelevant, yet the British election campaign will almost certainly throw up issues of such general importance as to be reported and even evaluated. In any case, there may be some general interest in the way in which the parties tackle the problem of how they would breathe new life into British research. But on the principle that single-issue politics are misguided, developments of this kind during the next few weeks will be reported objectively and, as far as possible, not discussed. An obvious exception to this rule would arise if the election should clarify the British (and French?) government's position on the Geneva arms negotiations. But there is much else to say and, mercifully, only three weeks of electioneering left. □

Shotgun divorce

The British government's reform of university finances is more hasty than wise.

How does a government spend taxpayers' money on good causes without inviting complaints of nepotism or giving licence to the exploitation of what is known in the United States as pork-barrel politics — the trade of parliamentary votes for constituency favours? The traditional technique is to leave the distribution of funds to a committee appointed for that purpose only. To function well, such a system requires that the committee should be independent of the political process. Ideally, it should also be independent of the recipients of the funds disbursed, a requirement often conflicting with the need to understand what recipients want. The British University Grants Committee (UGC), originally known as a commission, has been lauded as a successful device of this kind for most of the present century, but it is now to be radically transformed, and in the process renamed the Universities' Funding Council (UFC). There is to be a similar cake-cutting council to cater for the polytechnics and the other institutions of higher education, which are being transferred from local education authorities to the central government.

Although the UGC has been widely imitated elsewhere, in India for example, there is much in the Croham committee's view, earlier this year, that the time has come for change. Part of the trouble is that the scale of spending (more than £1,800 million a year) is too great to be managed equitably by a committee of amateurs, mostly academic, however enlightened. That is why the British government has plumped, as an alternative, for a small group of people functioning more or less as a board of directors of what might be called Universities plc (where 'plc' is the newish designation of a limited company). The analogy is made apt by the government's determination that the new council should be somehow close to 'industry' and by its announcement, a few weeks ago, that funds would be channelled to polytechnics and other institutions in the public sector only after formal educational contracts had been negotiated with them. The position of the universities was, at the time, undefined. Now it seems that the same principle will apply.

The draft proposals now published (*Nature* 327, 92; 1987) are either cosmetic devices or dangerous. The intention is that the Universities' Funding Council should negotiate for the provisions of educational services with either entire institutions, consortia thereof or, in suitable circumstances, individual university departments. In the language of the management schools, the

device is an excellent means of making sure that universities are publicly accountable; institutions failing to live up to their promises will presumably fall off the list of those from whom future tenders are invited. But the system is obviously also a means by which the character of higher education may be arbitrarily transformed by the attachment to the proposed contracts of conditions inconsistent with good academic practice. There is nothing in the new rules to prevent the proposed council imposing arbitrary deadlines for the completion of graduate courses or requiring that prescribed proportions of, say, engineering courses should be taught by people with industrial experience (which might be impractical).

Freedom

It will be easier to tell whether UFC will avoid these obvious pitfalls when the composition of the first council has been decided. Then at least it will be possible more accurately than at present to guess at the government's intentions. But it is hard to see how even the most enlightened application of the new principles can fail to undermine the autonomy of the universities. Academic freedom is not a licence for self-indulgence but rather, the freedom to decide 'what to teach, to whom and by whom'. If the system for the future will require that universities should go cap in hand to UFC each time they wish to introduce a new course of instruction, their freedom to teach what they consider to be in the best interests of their students will be lost, as will be the potential of the publicly supported system of higher education for academic innovation. The notion that individual departments might win educational contracts off their own bats is even more dangerous, suggesting a means by which the collegiate cohesion of universities will be destroyed. That is why it must be hoped that the effect of the changes planned will be indistinguishable from the present system, meaning that the change is nothing more than trendy semantic wizardry.

The Committee of Vice-Chancellors and Principals was right to demand last week that there should be an opportunity for genuine consultation on the government's proposals. It might have gone further and have complained that no meaningful application of them can be free from great dangers to the integrity of the present system. It would also have been proper to complain that the time allowed for consultation is far too short; the universities have until the end of June to comment on the details of the new proposals. The government's intention (on the assumption that it is re-elected on 11 June) is plainly that the present system should be changed in time for the academic year beginning on 1 August 1988. Universities will no doubt be expected to convert the academic plans on which they have been slaving into their first tenders for contracts from UFC.

The government's impatience is mystifying, and may prove to be self-defeating. The university system it so evidently distrusts is now changing more quickly than ever. There is a danger that plans for radical change, typified by the recommendations of the Oxburgh report (to UGC) for the Earth sciences, will be cast aside in the scramble for contracts with UFC. A greater danger is that similar patterns will be arbitrarily enforced by UFC, further sapping the self-confidence of the British universities. Yet the upheaval now proposed may not — and should not — last very long. If the system of British higher education is to suffer a further trauma, it would be best that it should take the form of a means of bridging the persisting gap between the universities and the rest of higher education. The government is unwilling to take this leap, fearing that the polytechnics would so quickly ape the universities that their apparently lower costs would be inflated. Yet the similarity of the two funding councils suggests that the government is hoping that some kind of common structure will emerge. Would it not be preferable to bite that bullet now, but more deliberately, allowing time for the new arrangements to be worked through in greater detail — and giving weary institutions a modest breathing space before the next upheaval is upon them? □

SDI facing new legal and political hurdles

- Congress votes to restrain SDI spending
- ABM treaty may outlaw even first stage

Washington

THE US Congress is asserting itself in its fight with President Ronald Reagan's administration over the Strategic Defense Initiative (SDI). Last week Congress moved forward on two fronts, cutting spending for research into anti-missile defences, and calling for compliance with the traditional interpretation of the Anti-Ballistic Missile (ABM) treaty which would hinder many tests of SDI technology.

The Democrat-controlled House of Representatives voted to cut the 1988 SDI budget to \$3,100 million, close to half of the \$5,700 million requested by Reagan, and \$500 million less than is being spent on SDI this year. A slowdown in the pace of SDI research is the last thing the administration wants. Despite the recent sober estimate of beam-weapon prospects from the American Physical Society (see *Nature* 326, 815; 1987), the administration is increasingly optimistic that technological problems can be overcome and a missile defence be created. Defense Secretary Caspar Weinberger now says that the first phase could be in place "as early as 1993 or 1994".

That first phase would not consist of 'beam' weapons but of more conventional 'kinetic kill' weapons — rockets and projectiles that would destroy missiles by high-speed collisions. Weinberger has urged the president to approve tests of interceptor rockets as soon as possible. He foresees eventual deployment of 3,000 small rockets stored in some 300 patrolling satellite 'garages' and capable of destroying 10 per cent of Soviet missiles. But the ABM treaty may yet stand in the way of tests.

The ABM Treaty, signed in 1972, specifically prohibits, in Article V, the development, testing or deployment of anti-ballistic-missile systems or components which are "mobile land-based, or sea-, air- or space-based". That would seem to mean that anything more than basic research on SDI technology is prohibited. But this is now dubbed the "narrow" view of the ABM treaty. The "broad" view, favoured by the administration, focuses on an ambiguity in "agreed statements" appended to the treaty, that could be taken to mean that "systems based on other physical principles" (than those known in 1972) are not covered in the ban. All but one of the key officials involved in the original treaty negotiations dispute

that this was what was intended. But a team of State Department lawyers, with access to the secret negotiation record, has just released a new report arguing for the "broad" interpretation.

While President Reagan has yet to commit himself finally to either view of the treaty, Democrats in Congress have been struggling to impose the narrow interpretation. The House has voted to forbid activities that violate that interpretation of the treaty unless warning is given to the Soviet Union that "supreme interests" have been threatened. Things have not gone so smoothly for Democrats in the Senate. Attempts to amend the 1988 Defense Authorisation Bill so that SDI development would be confined to the narrow interpretation, unless a joint resolution from Congress decrees otherwise, has yet to penetrate a cleverly constructed Republican defence shield preventing the matter being brought to a vote. Senator Sam Nunn (Democrat, Georgia), chairman of the Senate Armed Services Committee, has vowed to keep on trying, however long it takes. Two further attempts to obtain a vote will be made in the coming week.

Even if the broad view of the ABM treaty prevails, SDI will still be in trouble, according to Nunn. The 'kinetic kill' vehicles in which Weinberger expresses faith as the first stage of SDI, would not seem to employ "other physical principles" but to be little more than sophisticated interceptor rockets, and in any case, are "mobile land-based" weapons specifically prohibited by the treaty. While the latter objection can be overcome by making the kinetic kill vehicle a fixed system for testing, the former will require some clever lawyers to find a way around.

Given sufficient determination from President Reagan, neither budgetary constraints nor legal arguments over the ABM treaty will slow SDI development. A final budget will be decided by both the Senate and House of Representatives and is likely to be much higher than the House figure. Constraints on treaty interpretation can be vetoed. But with Reagan in difficulties over his administration's involvement in funding Nicaraguan rebel groups, Congress may come to have a greater say in what happens in SDI than it has in the past, and that may mean no commitment to rapid deployment.

Alun Anderson

DoE hopes for superconductor cooperation

Germantown, Maryland

RESEARCH by US industry into the new superconductors is gaining momentum, judging by attendance at a meeting at Department of Energy (DoE) headquarters last week. The two-day event was organized to review the efforts of DoE laboratories on the new materials, and to give industrial researchers a chance to show off their own work. Although there is no sign yet of a coordinated national research programme, the DoE is trying to offset worries about Japanese progress by increasing its visibility as a broker for collaboration between universities, national laboratories and industry.

High current densities, recently achieved in thin films grown at IBM's Yorktown Heights laboratory, provided the major scientific interest. Because the ceramics have strongly anisotropic electrical properties, David Clark and his colleagues grew single grains, about a micrometre thick and up to a centimetre long, so that the crystal plane in which electron pairs travel was parallel to the film. The current density achieved across the planes was as much as seventy times smaller.

Panels of industry and university scientists gave their views on research progress and on prospects for superconductor applications. There was a general note of optimism, despite the acknowledged difficulties in making use of the fragile ceramics. Many participants expressed worries about the chemical as well as the mechanical durability of the materials; the oxides absorb water and carbon dioxide readily, changing the oxidation state of the copper atoms and destroying superconduction. But any difficulty was no sooner mentioned than countered instantly with suggestions for remedies. As panellists emphasized, many fairly obvious and straightforward experiments are waiting to be done, as soon as researchers can find the time.

DoE seems to see its purpose as coordination, not direction. Participants generally agreed that industry will find any research applications under its own steam, and that the role the universities and laboratories could best fulfil is to come up with new materials and theoretical underpinning for their properties. Angela Stacy, of Lawrence Berkeley Laboratory, lamented the lack of strength in solid-state chemistry in the United States; gas phase chemistry and reaction dynamics are perceived as being more intellectually exciting. But the idea that DoE might try specifically to encourage research in certain areas was not enthusiastically received by DoE officials at the meeting.

David Lindley

Plant breeding is the last shot in privatization locker

London

AFTER several months' delay, the British government last week put up for sale its institutes for the development and marketing of new varieties of plants. The plant breeding assets of the Plant Breeding Institute (PBI) and the National Seed Development Organisation (NSDO) are to be sold as a single package to the highest bidder who can satisfy the government that it is "committed to the continuation of plant breeding relevant to UK agriculture and horticulture". But the delay, blamed on "special considerations" of splitting PBI, could have jeopardized the government's plans. The major opposition parties, Labour and the SDP/Liberal Alliance, said last week that if they were in power after the general election on 11 June they would halt the sale.

If the sale does go ahead, PBI's plant science programmes will become part of a new Institute of Plant Science Research to be established by the Agricultural and Food Research Council (AFRC). The NSDO workforce of 70 and half PBI's staff of 250 will transfer to the private

company. PBI's rapidly expanding oilseed rape programme will not be included in the deal, but will go to the Agricultural Genetics Company (AGC) in settlement of various claims AGC has on the institute. AGC, a private company, was set up in 1983 and given the first option to exploit plant biotechnology discoveries made by AFRC. AFRC will receive royalties on any new varieties produced.

The announcement last July of the government's plans to sell PBI was roundly condemned by the staff, who felt that the special link between the breeders and molecular biologists would be broken (see *Nature* 326, 822; 1987).

More than 80 per cent of NSDO's revenue is derived from royalties on varieties produced by PBI. In 1985–86, NSDO earned the government £6 million in royalties and profits. Given that PBI operated during the same period on an AFRC grant of £3 million, with a further £1 million from other sources, PBI fails to see why it cannot be allowed simply to finance itself. It assumes it is being sacrificed for the dogma of privatization.

PBI staff are unhappy at the way the sale is being handled by government-appointed merchant bankers Lazard Brothers. Potential buyers will be issued with a confidential memorandum of information; on the original timetable, a short-list will be drawn up in July, with the final purchaser selected by August, and the sale completed by September. Until the plant science programmes re-locate in March 1990, the private and public sectors will work on the same site. Offices and laboratory space will be moved as far apart as possible; the locks on the greenhouses will be changed — one kind for the private sector breeders, another for the public sector scientists. It is this kind of enforced segregation that has left many staff feeling bitter. "This is one of the most difficult and damaging aspects of the whole thing. For 75 years we have been trying to make science and breeding work together profitably and now we have to drag them apart", a senior staff member said. Staff are also unhappy at not being allowed to see the memorandum of information.

The government has refused to indicate the sort of price it expects the package, including 400 acres of land, to fetch. Unofficial estimates vary from £20 million to £70 million. Imperial Chemical Industries (ICI) and AGC have expressed an interest, and a four-man management team from NSDO and PBI will be hoping to put together a management buy-out, but their chances of success must be slim.

Simon Hadlington

Name change to recognize Levich's contribution

New York

Dr Benjamin Levich, once the highest ranking Jewish refusenik scientist in the Soviet Union, and latterly a holder of one of the five Albert Einstein professorships at the City University of New York (CUNY), will be permanently commemorated at New York City College — the CUNY campus where he taught in his last years. City College's Institute of Physico-Chemical Hydrodynamics, which was established in 1979, when Levich came to New York, and which has from the beginning been known informally as "Levich's institute", is now to bear his name officially. This change of name was announced last week at a memorial colloquium for Levich, who died earlier this year.

The "Benjamin G. Levich Institute of Physico-Chemical Hydrodynamics" will continue working in the field pioneered by Levich, which relates the physical properties of a flow to the chemical processes taking place in it. As the papers presented at the colloquium showed, Levich's research sparked interest in aspects of the subject ranging from helicity on vortex flows to the transport phenomena of two-phase systems and to the reflection of pulses from random media — a subject with significance for seismic surveying.

Just 10 years ago, far from commemorating Levich's name in the field for which he had done so much for Soviet science, the Moscow authorities set about removing his name from the literature altogether. Not only were his books withdrawn from circulation, but references to his work were excised from Soviet publications — and his name was blanked out from articles in Western journals before photocopies were distributed. This unusually severe reaction was probably because of his position as a corresponding member of the Soviet Academy of Sciences.

In 1977, when Levich was 65 — a birthday which in normal circumstances would have been honoured by a special conference in the Soviet Union — his colleagues and supporters abroad decided to hold a birthday conference *in absentia* in Oxford. This proved so successful, from the scientific point of view, that it was decided to hold regular biennial Levich Conferences in physico-chemical hydrodynamics.

When he was eventually granted permission to emigrate, Levich accepted a professorship at Tel Aviv University, Levich supervised research both in Tel Aviv and at City College, and the link between the two institutions will continue at least for a time, as his son, Dr Evgenii Levich, also has academic commitments to both.

Vera Rich

New UK plant lab

London

TWYFORD Plant Laboratories, the horticultural products company based in Somerset, England, is to invest £6 million in a new plant biotechnology laboratory to be built in the Cambridge Science Park. The 35 or so scientists at the new facility will undertake fundamental research in the molecular plant sciences with the aim of developing new products and services for the horticultural sector. Specific research areas will include improved diagnostic methods for plant disease, genetic engineering of plants to make them resistant to insect and pathogen attack, and the creation of novelty variation in ornamentals.

Work on the new building, with 1,300 square metres of laboratory space and a sophisticated greenhouse, is expected to start in June, and the facility should open early in 1988. Since 1982, Twyford has specialized in the supply of plants propagated by tissue-culture methods to growers worldwide, as well as providing plant health services. Numbers employed at their Somerset headquarters have increased by over 200 in the past four years, and turnover now exceeds US\$12 million. In siting the new laboratories in Cambridge, the company hopes to benefit from collaboration with local centres of excellence in molecular biology and in plant breeding and genetic manipulation.

Giles Courtice

Row over government support for West German universities

Munich

A row over the level of government support for university research in West Germany has broken out between the university rectors' association (Westdeutschen Rektorenkonferenz) and the research ministry.

At the association's annual meeting in Göttingen last week, the outgoing president, Theodor Berchem, appealed for more government support. He protested that the university research budget, estimated at DM 59,300 for 1987, is less than one-seventh of the funds currently spent on research by industry and government.

Research Minister Heinz Riesenhuber was quick to reply. He said that the biggest problem is quality, not quantity: "the money is there, but (good) people to use it are in short supply".

Riesenhuber's snappy reply provides a clue to several serious difficulties confronting West German university researchers, although few would agree that quantity is any less of a problem than quality. Universities have been filled far beyond capacity for at least ten years, reducing the amount of time professors can devote to their research. According to one rector, Gisbert Freiherr zu Putlitz at the University of Heidelberg, almost a whole generation of professors has done little more than teach, to the point where they are hardly capable of redirecting their efforts to research.

New academic positions are in short supply. Of every 1,000 people who achieve the level of *Habilitation*, a pre-professional post reached by the best young researchers, only seven have a chance at tenure-track positions.

The situation in West Germany is complicated by the large number of sources of funding for the universities. It is the responsibility of the Federal Ministry of Education (BMBW) to provide what federal support there is for the universities, which are also funded in part by the states. The Research Ministry (BMFT) provides funds for applied research in universities and *Hochschulen* as well as for space and nuclear research and other large-scale programmes. Grants for basic research issue from the Deutsche Forschungsgemeinschaft (DFG). All of these bodies, as well as state governments and industry, will have to play a role if the pressure on West German universities is to be relieved.

Maintaining the current level of quality in research would not be too hard, said zu Putlitz, as long as more money became available. "DM4 million would make a very big difference", he said.

Sometimes experiments cannot be per-

formed simply because there is not enough money for the electricity bill. The 2-3 per cent yearly increase in university budgets just barely covers personnel costs, said zu Putlitz. Furthermore, the university budgets are not as high as they look, since included in them are mandatory 5 per cent reductions which must not be spent by the universities.

To improve university research, rather than just maintain its level, more would have to be done. Zu Putlitz stressed the need to reduce the student-faculty ratio and hire young professors who are more research-orientated than previously.

Even if working conditions for researchers are improved, it does not mean that the best people will stay in West Germany for postdoctoral or junior faculty positions. The Heisenberg Fellows programme, supported by the DFG, and the Fiebigler Program both attempt to hang onto these young researchers for a few years until more tenure-track positions are available. But currently only 230 Heisenberg fellows are selected each year and the Fiebigler Program also supports relatively few people, and only in the wealthier states at that. According to Hubert Markl, president of the DFG, if the situation is not improved soon, it will have a "fatally discouraging effect" on young researchers in West Germany during the 1990s and beyond.

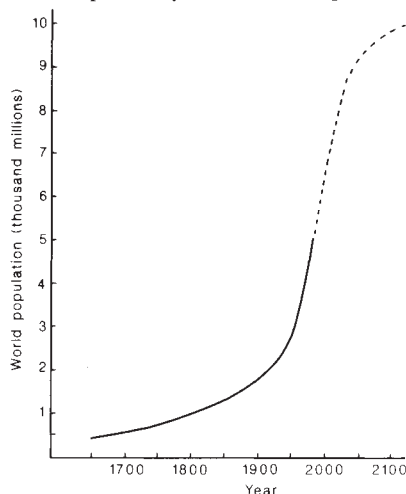
Steven Dickman

World population will soon reach five thousand million

London

THE world's population will shortly reach 5,000 million, according to the United Nations Fund for Population Activities (UNFPA). Every minute it grows by 150 people, every day by 220,000; 90 per cent of this growth is in developing countries.

An increase in growth rates started in the eighteenth century and reached its peak in 1970. According to the United Nations, the growth rate is now declining, and will probably reach zero again in a



From about 500 million in the middle of the seventeenth century, the world population has risen to 1,000 million in the early nineteenth century and to 5,000 million in 1987. The assumptions made in the UN report suggest a population of 6,000 million in 1999.

century from now, when the world population is expected to stabilize at about 10,000 million.

"Developing countries are now undertaking family planning", said Dr Nafis Sadik, UNFPA executive director. "The ten countries with the largest populations

are launching control programmes in a very active way. Even the countries of sub-Saharan Africa, which were never interested in population control before, have now adopted programmes for family planning. As a result, population growth levels will come down dramatically."

The declining growth rate for the developed countries has brought some sharp contrasts. Europe, including the Soviet Union, passed the half billion mark (500 million) before 1950; Africa attained it only in 1982. But Africa is expected to reach 1,000 million some time between 2005 and 2010; Europe will probably never reach it.

Forty per cent of the world's population now lives in urban areas; most of the world's largest cities are now in the developing regions.

In theory, according to the UNFPA, there is enough food for everyone, but its report quotes World Bank estimates that about one in six people worldwide do not eat sufficient calories for an active working life. In Africa, food production is up 20 per cent over the last decade, but *per capita* food production has fallen 11 per cent.

The State of the World Population report examines, and rejects, various arguments in favour of increased population growth. Sheer numbers of people, it concludes, do not contribute to economic development by increasing the potential for innovation. High fertility rates can gobble up many of the fruits of economic expansion, Sadik said, comparing Brazil to Japan.

The report also counters the idea that population growth does not affect the balance between humans and nature, pointing to deforestation and species extinction.

Kathy Johnston

British withdrawal from CERN again a growing possibility

London

BRITISH scientists working at the European Organisation for Nuclear Research (CERN) have formed a group to lobby the media, industry and politicians because of growing concern that Britain will pull out of the venture. The ten-man CERN United Kingdom working party, chaired by Dr Eifionydd Jones, says that withdrawal from CERN would be "an act of scientific vandalism". The group is preparing to launch a concerted campaign to impress upon science policy-makers and industry the existing and potential importance of CERN. The working party argues that compared with many of the other member states, Britain's science community benefits to a proportionally greater extent, and that accessibility to CERN's facilities "acts as a plug to the brain drain". The group is also concerned with the apparent indifference towards CERN shown by British industry, whose failure to bid competitively for high-technology contracts is seen as one reason why Britain's future with CERN is uncertain.

Britain's continuing collaboration was put in doubt with the publication in 1985 of a report by a panel chaired by Sir John Kendrew that recommended that the United Kingdom should remain a member of CERN until 1989, but should continue its membership beyond that time only if it could be achieved at a significantly lower cost. It recommended progressive reduc-

tions in subscriptions of about 5 per cent in 1988-89, rising to a total of 25 per cent in 1991-92. The report was endorsed by the Science and Engineering Research Council (SERC) and the Advisory Board for the Research Councils. In 1985-86, Britain's contribution was £37.9 million. This increased to £55 million in 1986-87 because of a sharp fall in the value of the pound against the Swiss franc, the currency in which the contributions from all 14 member states is assessed. In response to the Kendrew report, Britain proposed, in February 1986, a seven-man international CERN review committee, which was set up under the chairmanship of Professor A. Abragam. The review committee will make an interim report to the CERN council committee on 18 June, when it will present its principal findings. It is unclear whether the committee of the council will act on any of the preliminary recommendations, or wait for the review committee's final report in December. It is felt that the latter is more likely. But, given that Britain requested the interim report, it is possible that a decision on Britain's future could be taken on the basis of the findings presented in June.

Under the rules for withdrawal, Britain would have to give notice before the end of the year if it wished to pull out before 1989. How long CERN council will take in deciding whether to act on any recommendations, and whether any action

taken will satisfy the British government is not known. It is generally felt that the level of reductions recommended by the Kendrew committee is unrealistic if CERN's scientific capability is not to be severely affected. Apart from the feeling within the physics community that Britain's withdrawal would be in itself detrimental, observers argue that it would destroy Britain's credibility as a serious international collaborator, something, they say, it can ill afford, particularly as it is seeking European partners for such projects as the ISIS pulsed-neutron facility in Oxfordshire. **Simon Hadlington**

US companies unite on chip

Washington & Tokyo

THE United States is to imitate Japan with an industry-wide, government-supported cooperative project to win back the world lead in semiconductor chip production. Japanese companies, long the top suppliers of standard memory chips, overtook the United States in total semiconductor sales last year.

Cooperative research projects, involving companies that usually compete, have been organized in Japan by the Ministry of International Trade and Industry (MITI), and have greatly helped Japanese companies to refine manufacturing techniques. But until recently they have been more than US companies could contemplate. Now, a coalition of US semiconductor companies, materials companies, equipment manufacturers and universities have got together to develop advanced semiconductor production equipment. The 'Sematech' project's aim is not to design new chips, an area in which US companies are still unsurpassed, but to do just as the Japanese did and refine manufacturing processes so that the maximum yield of defect-free chips can be gained.

Financial backing for the project, at \$250 million a year, is being sought in equal parts from industry and from government. Appropriations for Sematech are already included in the Department of Defense's 1988 budget proposals, now on their way through Congress. Charles Sprock, president of the National Semiconductor Corporation, who is leading the project, said he believed the United States could draw level with Japan in five years and "regain the lead soon after".

In Japan the formation of Sematech has been welcomed. "Playing Japan's game" and producing goods of high enough quality and low enough price to sell in Japan is seen as a very welcome alternative to a trade war that denies foreign markets to Japanese products.

Alun Anderson & David Swinbanks

British semiconductor company enters Japan's back door

London

WHILE most of the major US and European semiconductor manufacturers are doing battle with the Japanese, claiming trade malpractice through dumping and demanding greater access to the Japanese market, Inmos, the small British chip company, is selling its products in abundance to Japan.

Such is the group's success in Japan that a new company, Inmos Japan KK, has been formed to exploit what it considers to be an expanding market because of interest shown in the Inmos transputer, a revolutionary microchip design concentrating memory, computational power and communication links on one device.

The international arm of Japan's telecommunications network, KDD (Kokusai Denshin Denwa), has launched the first of the country's designs based on the Inmos device, an image-processing system, developed in conjunction with the image-processing equipment manufacturer

Kashiwagi Laboratories.

Japanese interest in the transputer has given hope to Inmos, formed in the late 1970s with government funds to spearhead Britain's entry into a new semiconductor age, that it could break even in this coming year. The group, which has an arm in Colorado Springs in the United States and two in Newport in Wales, has been dogged with controversy, and financial problems. The project consumed more than £65 million in government funds and loans since it began and it was sold, in the autumn of 1984, to the British electronics group Thorn-EMI for £95 million.

The group then had to attempt to survive one of the worst recessions in the global semiconductor market. The group has been rationalized, retaining a development and design presence in the United States and about 400 people, and design and management expertise in Bristol and manufacturing in Newport.

Bill Johnstone

West German students demonstrate against reforms

Hamburg

WEST German students are on the streets, demonstrating against plans, in states controlled by the Christian Democratic Party (CDU), to reform education in the Gymnasium (secondary) schools. CDU Kultusministers (in charge of Culture, Education and Church Affairs) in these states believe that the *Oberstufenreform* (reform of the upper schools) has failed, but in making their case they present the same facts that were used 15 years ago when all Kultusministers accepted a change from prescribed classes to a course system with a more or less free choice of subjects. At that time it was claimed that students entering university were inadequately educated in specialized subjects.

The upper school system was reformed to allow students to choose the subjects most appropriate for their university career, but many students reacted by choosing just the easiest courses. One reason for this is that the university entrance requirements for many subjects call for a good average school performance without specifying particular subjects. According to the heads of the universities, entering students are still poorly prepared in certain subjects. The

CDU Kultusministers now say that it is no longer acceptable that students can leave the gymnasium without having studied essential subjects, particularly German and mathematics, in their final three years. To remedy this lack of a comprehensive education, the Christian Democrats now want to reform the reform by requiring study of German, mathematics, a foreign language, one natural science and one fine arts subject, and sports.

Another crucial question is whether to include vocational education in the gymnasium curriculum, as is now done in the *Kolleg-Schulen* in Nordrhein-Westfalen (NRW). The CDU wants to eliminate such secondary schools altogether. However, the Social Democratic Party (SPD), which controls NRW, Saarland, Bremen and Hamburg, sees vocational schools as models for the future. More than 25,000 students have demonstrated in Hamburg during the past few weeks, and across the country more than 120,000 students have called strikes and even occupied their schools. Widespread opposition to what the students call *Abitur-Deform* has forced the Kultusministers to hold a special meeting on the subject during their next conference in June. **Jürgen Neffe**

Rocket failure identified

Washington

THE Atlas-Centaur rocket that had to be destroyed by ground controllers less than a minute after launch on 26 March (*Nature* 326, 428; 1987) should never have left the ground, according to the findings of the board investigating the accident.

NASA (National Aeronautics and Space Administration) investigators found that the rocket had been launched into "atmospheric conditions conducive to triggered lightning", in violation of safety rules. A single lightning strike was the most probable cause of the accident although it is still not known how the lightning reached the digital computer unit after striking the nose fairing. The surge of electricity changed a single word in the computer memory essential to control yaw of the rocket. The board will make recommendations to tighten the weather criteria used to decide whether a launch can go ahead.

Alun Anderson

Call for destruction of smallpox virus

London

A WORLD Health Organisation (WHO) committee is asking for remaining stocks of smallpox virus stored in Atlanta and Moscow to be destroyed later this year, and

countries that are still vaccinating their armed forces against the smallpox disease are being urged to stop.

"There is no scientific reason for keeping the virus stocks", said Professor Arie Zuckerman, the British government's chief adviser on vaccines. The genetic structure is available because the genus has been cloned, he explained. "It's a virulent virus, and there is always some risk it can escape."

Several countries, including the United States and the Soviet Union, still immunize their military recruits against the now-eradicated smallpox disease. "We are concerned at this military immunization", said Dr Z. Jezek, WHO's smallpox expert. "It is apparently due to their fear of biological warfare."

But medical sources said smallpox would be a useless tool of biological warfare because it does not spread easily and an effective vaccine can be applied quickly.

The vaccinia vaccine can carry the risk of side-effects, including brain damage. In addition, vaccinia can trigger the disease in someone who has been infected by the AIDS virus.

Kathy Johnston

Erratum

Marc Aaronson, who died on 30 April (see *Nature* 327, 92; 1987) was born in 1950. □

Soviet payload comes down in Pacific Ocean

London

THE Soviet Union's 'new generation' space launch vehicle *Energiya* was tested last weekend with a payload of 100 tonnes. The rocket performed as planned, but the payload, described as a full-size-and-weight mock up of an orbital craft, failed to go into orbit and came down in the Pacific.

The *Energiya*, according to the Soviet press agency TASS is now "the most powerful rocket in the world". This claim seems justified; the maximum NASA (National Aeronautics and Space Administration) payload so far is the 30-tonne space shuttle. Following the policy of *glasnost*, the launch of *Energiya* was shown on Soviet television, but only the following day. Several recent Soviet launches have been shown live on television, but these were of *Proton* rockets, the work-horses of the space programme.

The *Energiya* will be used both for reusable and expendable spacecraft. It is only since Mr Mikhail Gorbachev came to power in 1985 that the space planners have admitted to working on reusable spacecraft, although Western Soviet-watchers were aware of a Soviet 'shuttle programme' several years earlier. In the pre-Gorbachev era, the Soviet publicists seized on the military involvement in the US-shuttle programme, and contrasted it with the Soviet *Salyut* programme, serviced by expendable *Soyuz* and *Progress* rockets, and which was, they stressed, entirely devoted to peaceful ends. Now, however, the Soviets have not only admitted to a 'reusable' programme, but have even implied that military experts were involved in the development of *Energiya*. Since, so far, there has been virtually no spin-off from Soviet military research to the civilian economy, this implies either that the *Energiya* rocket (and by implication, the Soviet reusable programme) has a military dimension, or else that the barrier between the military and civilian sectors is becoming less rigid. During his visit to Baikonur last week, Gorbachev was accompanied by the Minister of Defence.

The unscheduled splash-down of the *Energiya*'s payload, although clearly a disappointment for all concerned, has been neatly turned by the Soviets into a moral, underscoring the current drive for greater efficiency and quality-control. "Check and weigh everything one more time!", Gorbachev reportedly told the *Energiya* team last week. Someone, it is implied, failed to heed his words.

Vera Rich

International harmony on plans for the US space station

Washington

THE prospects for international collaboration in the US space station brightened last week after two days of discussion in Washington between representatives of the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA).

Andrew Stofan, associate administrator of the NASA Office of Space Station, said that after "eleven months of struggling, everything seems to have broken free at one time", both because of progress in international negotiations and because of support given to space station funding in Congress.

International collaboration had seemed on the verge of collapse after US Defense Secretary, Caspar Weinberger, had insisted that the Department of Defense be allowed to use the station for military research (see *Nature* 326, 727; 1987), in defiance of earlier promises that the station would be used for peaceful purposes only. Now there has been progress towards a form of words for an agreement

that satisfies European desires that the station be used for peaceful purposes consistent with international law, while leaving certain military options open.

Negotiations in Washington were not disturbed by a row between Weinberger and Representative Norman Mineta (Democrat, California) who had proposed a bill to prohibit use of the space station for weapons research. Assurances were given that Weinberger's angry protest comparing Mineta's ideas to "Soviet propaganda" was not aimed at an international audience.

ESA representatives are also very pleased, according to Stofan, that NASA has accepted the man-tended free-flying laboratory (MTFL) as part of the project. MTFL will be launched by an Ariane 5 rocket and rendezvous with the space station twice a year for refurbishment. The rest of the time it will fly free, providing a low-gravity environment for experiments, undisturbed by dockings of the space shuttle and movements of space station personnel.

Despite NASA's official optimism, there are still risks that funding will not be found, and that the project will not survive changes in the administration that will follow next year's presidential election. Stofan hoped that an independent review of the station's costs and technical plans now being carried out by the National Research Council would help provide a "bridge from one administration to another".

Negotiations with Japanese representatives, who have consistently opposed military use of the space station, will take place next week. Satisfactory talks have already been held with the Canadians, the fourth partner in the station.

Alun Anderson

Medical professors' public complaint

London

IN an unprecedented public statement, the Association of Clinical Professors of Medicine has highlighted the demise of non-drug-related medical research and the "rock-bottom morale" of research staff in Britain.

The criticisms, which are levelled at the government's attitude to the support of research, are contained in the association's submission to the subcommittee on priorities in medical research of the House of Lords Select Committee on Science and Technology. The association, whose membership is composed of 150 university-based professors of medicine and related clinical subjects, for the first time has "now expressed publicly deep anxieties about the future of medical research in the UK". While the association admits it is experiencing what scientists in other areas have already identified, it claims that there are problems resulting from the present government's policies "which affect clinical research in particular".

The statement to the House of Lords claims that the loss of certain university funds is having a dramatic impact on the direction of research effort. Greater reliance is now being placed on the pharmaceutical industry, which provides money for clinical trials of drugs. As a consequence, substantial areas of medical research are being neglected as "industry is not in general prepared to invest in non-drug-related research, such as preventive medicine, epidemiology, mental illness or the problems of ageing". The association, whose members hold chairs in departments of medicine or allied specialties, blames government cutbacks. Good medical research depends on first-class scientists and good projects. "These we have, until now, had in abundance, though we are now losing more and more good people abroad."

Bill Johnstone

Government money on the table for bankrupt college

London

THE British government has handed a £10 million lifeline to University College, Cardiff, to rescue it from possible bankruptcy, after charges of serious financial mismanagement. The interest-free loan and grants will be made only on the condition that the college dismisses or obtains the resignation of its principal, Dr C.W.L. Bevan, and "takes all reasonable steps" to sue or surcharge any staff responsible for financial losses to the university.

The rescue package is spelled out in a letter from the University Grants Committee (UGC), after months of investigation into the college's financial crisis. The government had earlier warned that the institution's grant would be withdrawn unless there were changes in the top management. Last week saw the resignation of four senior staff, and the bursar resigned earlier this term.

The college was facing debts of up to £7 million, brought about by what an inquiry set up by the acting principal Professor Lee Sheridan called a "tangle of uncosted schemes, arbitrary authorizations and irregular salary payments". The £10 million will be made available only if University College, Cardiff, amalgamates with the University of Wales Institute of Science and Technology, and overhauls its financial management.

Academic and financial plans for the new amalgamated institution must be approved by the UGC. A meeting is expected soon between the college and the UGC to discuss details, including proposals made by a college committee which would mean nearly 140 redundancies and the amalgamation of some departments. "The money will give us breathing space, but it doesn't solve the problems", a college spokesman said. The rescue package includes a £5 million loan from the government, which will have to be paid back by the sale of buildings and land when the new institution is set up; other funds in the package will cover staff redundancies.

The college's principal, Dr Bevan, has been on sabbatical since March, and he has not yet said if he will resign. He has been principal of the 6,000-student college for 21 years, and has survived a number of no-confidence motions by the college senate and its staff. Last month the college council refused to dismiss Bevan despite a recommendation from an inquiry into the financial crisis.

The Department of Education and Science has been scrutinizing the finances since early last year, when it commissioned an independent report which criticized the management. Bevan was quoted as calling the report "a bag of wind".

Kathy Johnston

Double British jeopardy for European laboratory's plans

Heidelberg

BRITAIN and the Netherlands, longstanding thorns in the side of the European Molecular Biology Laboratory (EMBL), are once again creating problems. Both countries have demanded a review of parts of the laboratory's scientific programme before deciding whether to agree to its plans for expansion, largely in biocomputing. And by remaining unwilling to agree even to the much-reduced research budget request of the European Commission (*Nature* 327, 5; 1987), Britain is putting the laboratory's biocomputing plans in double jeopardy.

The expansion in biocomputing, an

enhanced programme of courses and workshops and further support for the Hamburg Outstation at the Deutsches Elektronen-Synchrotron Laboratory are EMBL's three main plans for 1987-90. The scientific programme that incorporates these plans has been approved by the 14 countries that finance EMBL, but Britain and the Netherlands have not approved the 2 per cent annual increase in budget needed to finance the expansion, which would cost Britain about £150,000 a year. Instead, they have delayed a decision while awaiting a review of current activities from the scientific advisory committee of the laboratory. Matters

should come to a head in July.

To accommodate the expanded biocomputing programme and to provide the auditorium necessary to increase its teaching activities, EMBL needs to add a £2.3-million building, which would be completed in September 1988 if the budget is approved. The biocomputing programme is budgeted at close to £3 million, of which just under a third should come from the Land of Baden-Württemberg, and just over a third from the European Economic Community's bioinformatic programme, once British objections to the European research budget can be resolved.

There will be two main strands to the expansion in biocomputing under Chris Sander. One involves increasing research on protein folding, with emphasis on improving the ability to predict a protein's secondary structure from its linear sequence of amino acids. The other is to strengthen EMBL's role as a European centre for data collection and analysis.

EMBL is already the European half of the (losing) battle to keep an up-to-date database of all nucleic-acid sequences. An

Cuts in defence research hit low morale of UK industry

London

A MODEST reduction, in real terms, in the United Kingdom's defence research and development budget for the coming year has provoked criticisms from the country's high-technology industries, which accuse the government of 'running down' Britain's research effort in favour of more procurement from overseas. The issue is particularly sensitive because defence-related projects account for more than half of total research spending in the United Kingdom.

The criticisms surfaced in the wake of the publication of the government's white paper (policy document) outlining expenditure on defence for the year. It showed the research and development budget with a small increase to £2,350 million from £2,260 million last year. The government was quick to justify the change, claiming that too few commercial products evolve from defence-related research, that more resources should be devoted to the civil sector and that no future support would be given to programmes that seem to duplicate those of Britain's NATO (North Atlantic Treaty Organization) allies.

The new policy was stoutly defended in the budget paper, which claims that the government "shares the underlying concern of those who fear that necessary investment in defence research and development may crowd out valuable investment in the civil sector".

The critics maintain, however, that the government has shown a marked reluctance to support civil research in Britain and in Europe, in collaboration with its partners in the European Economic Community, while paying little heed to

the growing numbers of British scientists emigrating to work in the United States.

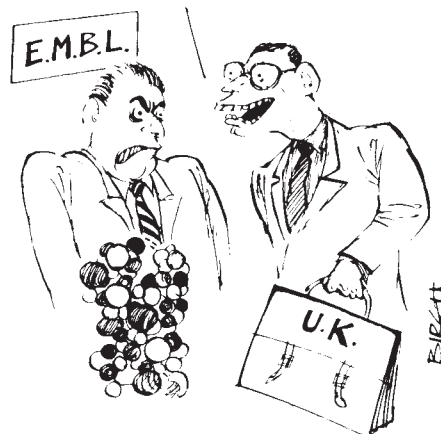
The UK defence industry has recently suffered two setbacks — the Royal Air Force rejected the GEC-developed Nimrod early-warning aircraft in favour of the Boeing AWACS on the grounds of radar performance, and Plessey's Ptarmigan battlefield-communications system lost the competition with a French system for a contract with the US marines. The report does little to allay fears that further disappointments may be in the offing, saying that special emphasis will be given to "avoiding duplication of successful equipment developments already achieved by our allies... our policies of increasing competition in procurement, and encouraging greater international collaboration to meet the equipment needs of the Alliance, are already aimed at ensuring more effective use of Britain's research and development resources".

Last year the government set up Defence Technology Enterprises Ltd (DTE), a company created to encourage the commercial exploitation of Britain's defence research. The budget paper refers to 15 licences that have been agreed or are reaching completion, through the DTE, for "exploitation of innovative technology from the research establishments". The report also updates the work of British researchers on the US Strategic Defense Initiative. By the end of last year, \$34-million worth of work had been awarded to 36 UK companies and academic institutions. Some 400 companies and 100 academic institutions have expressed an interest in conducting research for the programme.

Bill Johnstone

Statement on the Defence Estimates (Cmnd 101-111, HMSO, £5.50).

Look on us as your personal
Termination Codon.



important part of Sander's plan is to enable 'cross-talk' between the nucleotide sequence database and related databases, in particular the protein sequence database of the National Biological Research Foundation in Washington DC and the protein structure database at Brookhaven National Laboratory.

If the biocomputing programme is a success, it will be a feather in the hat of EMBL's director-general, Lennart Philipson, whose initial five-year term has been followed by a three-year renewal which may yet be extended by two years to 1992. His biggest headache remains the inability of the laboratory to attract and keep the staff necessary to maintain a strong programme in structural biology. That, and the recent loss of the laboratory's team of gene mappers, under Hans Lehrach, seems to leave the laboratory incongruously weak in molecular biology for the time being.

Peter Newmark

Future of US space programme

SIR—Your leading article, "Who wants the US space station?" (*Nature* 325, 745: 1987), raises important questions which must be answered quickly by planners and participants in the US space programme. Three years of study of space station plans by responsible representatives of the United States and the international space science community have, in fact, identified important scientific uses for a large, manned science facility in low Earth orbit. These include support for major astronomical, solar-terrestrial and Earth observing projects as well as initiation of serious study of materials science and technology and life science in microgravity. However, the extent to which the current NASA (National Aeronautics and Space Administration) plans for space science and the space station are adequate and can be justified in the face of recent developments spawned by the Challenger accident is open to debate.

My personal concern arises from the following. First, for the foreseeable future, the space shuttle must be regarded as a fragile, irregular transport vehicle which will be hostage to a variety of practical considerations, including substantial oversubscription by its users. Even its flight frequency will remain in doubt for some extended period. Nevertheless, the current space station plans rely exclusively upon the availability of this vehicle for all aspects of station construction and supply. With an estimated flight rate of about 12 per year, the current space station plan requires the equivalent of 2.5 years of dedicated shuttle flights. Is this a responsible plan for such an important programme? Is this the only possible alternative?

Second, as a consequence of the Challenger accident, serious disruptions have arisen in the national space science programme. NASA's science plans before the accident included the equivalent of 76 shuttle payloads (on the shuttle and other unmanned vehicles) between fiscal year 1986 and fiscal year 1995. As a consequence, many proposed and funded experiments have been put aside, creating a crisis in space science that has yet to be addressed in any meaningful way. Unfortunately, as discussed below, the space station plans have evolved independently of this situation.

Third, estimates of the acquisition and operating costs of the currently planned space station have climbed because of new costing procedures, reduced space shuttle flight capabilities and new and improved

knowledge of space station system requirements. This has introduced a substantial delay in the projected start of the construction project and lengthened its duration. With the present plan, a moderate level of science work with the space station seems improbable until late 1995 or early 1996. This delay in the availability of the space station, taken in concert with the decline in science flights associated with the space shuttle, presages a decade of reduced opportunity for space science investigations, even taking into account the flight of major facilities such as the Hubble Space Telescope, Galileo, the Gamma Ray Observatory, the Upper Atmosphere Research Satellite and others.

Finally, one of the great difficulties arising from current NASA plans is the lack of opportunity to press on with important work in microgravity science and technology. Materials research in space is a new endeavour now being actively pursued by the Soviet Union and its partners. Spacelabs that could have helped the United States and its international partners bridge the gap to space station are now greatly restricted at a time when space station utilization is drifting later and later. Even the concept of man-tended pressurized modules, which could conceivably be provided by private industry, seems to be in disfavour. In parallel with the problems facing materials science, I also note that with the current space station plan, the study of long-duration effects of space flight on humans will be seriously delayed, and this must certainly have a negative impact on the speed at which we can even seriously consider manned planetary exploration at the turn of the century.

Based on the foregoing, it is apparent that serious thought should be given to altering present space transportation and space station plans to suit the needs of space users. Alternatives offered by new vehicles such as the so-called Heavy Launch Vehicle or new types of supporting space facilities need to be considered. Ideas that call for an initial scaled-down space station which can service science users substantially earlier than is now planned must be given close attention.

The answer to your question of "Who wants the US space station?" is that science users want a space station but not necessarily the space station that has been put forward in recent months. The traditional areas of space science and the new fields related to space microgravity are not well served by current plans for a space station which arrives too late and is based on a space transportation system which is unable comfortably to meet the needs of its users. To arrive at the right space sta-

tion, NASA must consider the serious plight facing its broadly based science disciplines and technology developers and come up with a more versatile plan which better satisfies user needs.

PETER M. BANKS

*STAR Laboratory,
Department of Electrical Engineering/
SEL,
Stanford University,
Stanford, California 94305, USA*

UK's bad example

SIR—As a scientist visiting Britain to improve my knowledge and further the international exchange of ideas, I had not previously realized how political pressures are undermining the science-base of the country. I wanted particularly to come to the Freshwater Biological Association (FBA), because it is one of the most respected of the world's environmental institutes.

I am concerned and astonished about what is happening here. A third of the staff are to lose their jobs. To accommodate the economic cut-backs by the government, important areas of science will be eliminated, including mycology, palaeolimnology and sedimentology.

Having experienced the consequences of similar policies in my own country — Brazil, where they were enacted for ideological rather than economic reasons — and seen how bad it is for the progress of civil science, I am astonished and disappointed that the basic sciences can be so disregarded in a developed country.

The run-down of Brazilian science caused several of our best scientists to leave the country so that they could continue their research programmes abroad. Some scientific areas almost completely disappeared and, decades later, some fields have not recovered. Now, in more enlightened times, we are rebuilding our institutes and universities, but only slowly, with great difficulty and at a very high cost.

This should not be allowed to happen in Britain. Other countries will take a dim view of its imposed decline. They will think that Britain has lost faith in its own abilities, and the reneging on scientific responsibility will be a poor example for the developing nations.

Economic problems will not be solved by dismissing scientists who have spent many years in training and gaining the experience to undertake their exacting work. The cost in lost morale cannot be counted.

F. A. R. BARBOSA
(Federal University of Minas
Gerais, Brazil)

*Freshwater Biological Association,
The Ferry House,
Far Sawrey,
Ambleside, Cumbria LA22 0LP, UK*

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

Making light-spots travel further

Careful shaping of the profile of a light beam makes it possible to increase greatly the distance over which the peak intensity is not degraded by diffraction.

ON the face of things, as schoolboys know, it is a kind of nonsense to think that it may be possible to arrange for a beam of light whose physical dimensions are finite to propagate through space without becoming broader and more diffuse because of diffraction. The usual and simplest explanation is based on Huygens's principle: each point on the instantaneous wavefront of a propagating beam can be thought of as the source of an outwardly propagating shell of radiation. For radiation propagating outwards from a point source, the wavefronts are spherical, and a sequence of ever-larger spherical wavefronts is obtained by the mutual interference of the hypothetical spherical wavelets travelling outwards from neighbouring points on each spherical surface. In the days when the lumeniferous aether was all the rage, this simple picture was one way of relating the supposed elastic properties of the aether to such quantities as the velocity of light and the electrodynamic constants of the vacuum. The same result obtains for plane waves travelling in one direction which are infinite in extent in the planes perpendicular to the direction of travel. All that is elementary. Equally, it is readily imagined that a plane wave which is not infinite in its lateral extent cannot travel without spreading. The Huygens wavelets near the middle of each wave front may interfere destructively except along another plane surface displaced slightly downstream, but the Huygens wavelets spreading from near the edges of the wavefront will not be neatly cancelled out by their neighbours except along the surface of the displaced plane, so that the travelling beam becomes fuzzy at the edges.

Of necessity, the spreading-out obtains simply as a consequence of the finite size of the propagating beam, and occurs even in a vacuum. The distance over which a beam of circular cross-section r will remain reasonably compact is measured by its cross-sectional area divided by the wavelength (which is why those who would use lasers in defences against ballistic missiles must think of building instruments with very large cross-section).

But now, it seems, there is a way round the problem. J. Durnin of the Institute of Optics at the University of Rochester has apparently managed to find a solution of the wave equation which represents a beam of light of finite cross-section which

is immune from this diffraction phenomenon. Moreover, Durnin and two colleagues at the same university, J.J. Miceli and J.H. Eberly, have been able to construct in the laboratory a realization of Durnin's wave (*Phys. Rev. Lett.* **58**, 1499; 1987). Maddeningly, Durnin's account of his discovery, which will no doubt eventually explain how he came by his solution, is cited as "in the press with *J. Opt. Soc. Am.*" (Physicists — and physics journals — are becoming terribly cavalier in this respect, chiefly under the influence of their huge trade in preprints. It seems nowadays quite common that articles based on pre-prints appear in print before the originals that inspired them. It is more unusual that an experiment to confirm the correctness of a theory should appear in advance of the explanation. But either way, the practice is confusing.)

Not unexpectedly, the beam is really rather special. The intensity is far from uniform across the cross-section of the beam but, rather, consists of a relatively intense central spot surrounded by an infinite series of annular rings of light whose intensity is inversely proportional to their radius (measured from the central spot). The fact that this description may be reminiscent of the diffraction plates by means of which the high-definition Schmidt sky-cameras are constructed is far from coincidental; in both cases, the trick is a solution of the wave equation which, in cylindrical coordinates, represents a beam travelling along the axial direction z ; the solution of interest is then the product of an oscillatory function of z and the Bessel function of the first kind of zeroth order, known in the trade as $J_0(ar)$, where a is a constant and r is the off-axis distance.

The most striking feature of this beam profile is that it represents a sharp central peak surrounded by a sequence of oscillations which tail off to zero inversely with the distance. The half-width of the central spot is inversely proportional to the constant a . It simply requires a little differentiation to show that the function is indeed a solution of the wave equation. But the fact that the oscillatory function of z is multiplied in the solution by a function of the radius only shows that the profile of the beam is literally unchanged on propagation, at least so long as the space in which the wave is travelling is for practical purposes infinite.

How to realize this state of affairs in the laboratory? What Durnin and his colleagues have done is to place an annular slit illuminated by well-collimated light in the focal plane of a lens of comparable dimensions. The slit must not be so narrow as to cause diffraction effects on its own account nor so wide as to confuse the lens. The output from the lens is a conical band of light. Direct measurement shows that the on-axis beam will propagate for roughly 1 m (under the conditions of the experiment) without substantial degradation of the central spot (whereafter it quickly falls to zero). Direct measurement has shown that a comparable spot-like beam with a gaussian profile decreases in intensity by an order of magnitude in a few centimetres.

The immediate application of this development is in those circumstances in which it is sought to make laser beams travel vast distances with undiminished intensity. It is worth noting that a perfectly collimated wave from a 2-cm diameter laser will have lost much of its central intensity after travelling for 1 km in a vacuum. Because the distance increases with the square of the radius, the modestly larger instruments have markedly better performance, but if it is in practice possible to enhance the distance over which the central intensity is not degraded by something like a factor of 100 by an appropriate shaping of the beam, many now impractical tasks will be possible (or present jobs can be done with smaller instruments). Curiously enough, there is no reference to the Strategic Defense Initiative among the references.

One curious problem remains. One of the other lessons learned at school is that the diffraction of a beam of finite cross-section is not merely an annoying consequence of Huygens' principle but a consequence of Heisenberg's uncertainty principle. But Durnin and his colleagues appear to have improved on that by a factor of a hundred or thereabouts. What can be the explanation? The authors provide the answer off their own bat: the explanation is that, although the central and practically important feature of the Bessel beam is its narrow central spot, the outlying annular rings of light intensity imply that the beam is, for practical purposes, infinite in its lateral dimensions. Both Huygens and Heisenberg survive.

John Maddox

Insect evolution

Molecules and morphology

Michael Akam

In the fruitfly *Drosophila*, the homoeotic genes of the Antennapedia and bithorax gene complexes (ANT-C and BX-C) specify the diversity of the different body segments^{1,2}. By studying the homologues of this gene family in other invertebrates, we can hope to relate major features of morphological evolution to changes in the genetic control of development. On page 247 of this issue³, Beeman reports one step in this direction. He shows that, in the beetle *Tribolium*, genes apparently homologous with both the ANT-C and the BX-C map to a single chromosomal cluster.

The family of homoeotic genes in the ANT-C and the BX-C probably derives from an ancestral *Antennapedia*-like gene that existed early in the course of metazoan evolution. Members of the family are still clearly related at the molecular level, by possession of closely conserved homoeobox domains¹, and of other small regions in the protein-coding sequence. Diversification within this gene family began before the separation of the lineages leading to vertebrates and insects⁵, but the precise set of segment-defining genes in *Drosophila* probably evolved in parallel with the increasing differentiation and specialization of segments during annelid and arthropod evolution.

Three principal homoeotic genes now lie within the BX-C of *Drosophila*: *Ultra-bithorax* (*Ubx*), *abdominal A* (*abd-A*) and *Abdominal B* (*Abd-B*). These control the diversity of the abdominal segments⁶. In addition, *Ubx* is involved in the most posterior thoracic segment (T3). At least three other segment-defining genes lie within the ANT-C — *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*) itself^{2,5,7}. These genes control the diversity of the mouthparts and thoracic segments. Each of these genes is now known to be expressed in a defined region along the antero-posterior axis of the *Drosophila* embryo^{8,9}, and each is necessary in that region for normal development of segment identity.

Homoeotic mutations have been described in a range of other insects. In *Tribolium*, mutations with homoeotic phenotypes define at least seven loci. Beeman³ in this issue reports the results of genetic mapping experiments showing that six of these seven loci are clustered on the second linkage group, within three map units. Moreover, the order of the loci on the chromosome is the same as the order of the segments which each mutation affects. At one end of the cluster, mutations affect the mouthparts; at the other end

they affect the posterior abdominal segments. This striking colinearity of gene position and morphological effect is observed also for genes within both the ANT-C and BX-C of *Drosophila*.

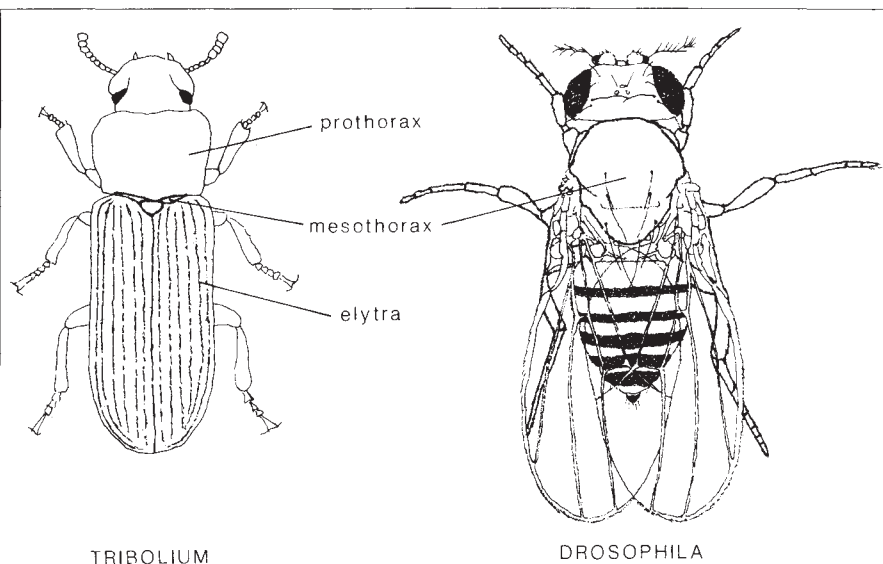
The simplest and most likely interpretation of these observations is, as Beeman suggests, that this single 'homoeotic complex' (HOM-C) of *Tribolium* is the functional and evolutionary homologue of both the ANT-C and BX-C of *Drosophila*, implying that the split into two separate complexes is a relatively late event in the evolution of the arthropods.

If this idea is correct, we might expect the homoeotic mutations that map within the HOM-C of *Tribolium* to have phenotypes closely analogous to those of *Drosophila*. This prediction is in part fulfilled, but there are intriguing differences. The mutation *alate prothorax* in *Tribolium*, for example, transforms the first thoracic segment (T1) to resemble the second, a transformation analogous to that caused by mutations in the *Scr* gene of *Drosophila*. In *Drosophila*, however, *Scr* mutations transform ventral but not dorsal structures of the adult. Even in homozygous clones of mutant cells, where the transformation might be expected to be viable, *Scr* mutations cannot make convincing notal struc-

tures appear in the adult prothorax⁷. In *Tribolium*, the dorsal transformation of the *alate prothorax* mutation is dramatic, resulting in the duplication of the elytral buds that give rise to the wing cases (see Beeman's Fig. 1d).

Another example concerns mutations in *Tribolium* that affect the development of the abdominal segments. One of these, *pas* (*pointed abdominal segment*), transforms A4, the fourth abdominal segment, to A3 — a transformation analogous to that produced by *iab-4* mutations in the *abd-A* gene of *Drosophila*. Another, *mas* (*missing abdominal segments*), is more puzzling. This mutation suppresses the development of A1, A2 and A3. No such mutation has been found in the bithorax complex of *Drosophila*, and I would expect such a lesion to be very rare. To achieve this effect would require the partial but specific inactivation of certain functions within both the *Ubx* and the *abd-A* genes, for in *Drosophila* *Ubx* controls T3 and A1, whereas *abd-A* controls A2, A3 and A4 (ref. 6).

How should we interpret such differences between *Drosophila* and *Tribolium*? They may reflect changes in the molecular organization of the genes which, although trivial in normal development, make some mutational events highly improbable in one organism, but relatively frequent events in another. At another level, they may result from differences in the way that the 'developmental mechanics' of flies and beetles interact with the segmental specification of cells to



The same or similar homoeotic genes probably exist in *Tribolium* and in *Drosophila* (see text). These insects belong to different orders, and probably last shared a common ancestor 300–400 million years ago. Both are relatively 'advanced' insects, undergoing metamorphosis during a pupal stage. The characteristic morphology of insects in these two orders is very different. For example, in flies (Diptera), the first and third thoracic segments are typically much reduced, and the single pair of wings are borne on T2, which dominates the anatomy of the thorax. In beetles (Coleoptera), T1 and T3 are well developed; T2 is characteristically rather small, but it bears the highly specialized elytra (wing cases) that hide the functional wings of T3 and the abdomen. It remains to be seen whether changes in the expression of homoeotic genes underlie these differences in development.

control the final expression of a homoeotic phenotype.

Such differences may also provide hints of more fundamental differences underlying the way that homoeotic genes control development in the two organisms. For example, if *mas* is a mutation affecting a single gene in the HOM-C of *Tribolium*, and if the transformation of A1, A2 and A3 is typical for such mutations, then the genetic control of the *Tribolium* abdomen cannot involve genes directly analogous to *Ubx* and *abd-A*.

It seems almost certain that fundamental changes in the functions of homoeotic genes must be involved in specifying the very different body plans of major groups within the annelid–arthropod lineage. It is much less clear whether such changes are relevant to the morphological evolution of different insect orders. The truth must lie between two extremes. Perhaps a constant set of homoeotic genes defines exactly the same segment identities in every insect embryo. The myriad and wonderful differences between insects would then depend on the differential interpretation of these conserved 'codes'. Alternatively, the profound modifications of segments characteristic of different insect orders may have been achieved at the very highest regulatory level, by modifying the domains and combinations of homoeotic gene expression.

A programme for further work is clear. First, to define, on the basis of molecular homology, whether the cast of characters (the set of segment-defining homoeotic genes) is the same in all insects and, indeed, to determine how it differs in different arthropods. Then, to investigate how each of these characters is deployed to specify the identity of segments. We know already that in *Drosophila*, the final segment pattern is the outcome of complicated spatial and temporal changes in the pattern of expression of homoeotic genes^{10,11}. Beeman's work with *Tribolium* may help to elucidate whether changes in these patterns contribute to the diversity of form among other insects. □

1. Lewis, E. B. *Nature* **276**, 565–570 (1978).

2. Kaufman, T. C. & Abbott, M. K. in *Molecular Aspects of Early Development* (eds Malacinski, G. M. & Klein, W. H.) 189–218 (Plenum, New York, 1984).

3. Beeman, R. *Nature* **327**, 247–249 (1987).

4. Gehring, W. J. in *ICN-UCLA Symposia on Molecular and Cellular Biology* (1986).

5. Regulska, M., McGinnis, N., Chadwick, R. & McGinnis, W. *EMBO J.* **6**, 767–777 (1987).

6. Sanchez-Herrero, E., Vernos, I., Marco, R. & Morata, G. *Nature* **313**, 108–113 (1986).

7. Struhl, G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7380–7384 (1982).

8. Harding, K., Wedeen, C., McGinnis, W. & Levine, M. *Science* **229**, 1236–1242 (1985).

9. Akam, M. *Development* (in the press).

10. Akam, M. & Martinez-Arias, A. *EMBO J.* **4**, 1689–1700 (1985).

11. Chadwick, R. & McGinnis, W. *EMBO J.* **6**, 779–789 (1987).

Michael Akam is in the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

Superconducting ceramics

Is there an explanation?

Nevill Mott

THE short answer is that there are many, perhaps as many as there are theorists active in the field. An indication of the kind of theory to be expected was recently given by the Editor of *Nature*¹. The others that I have seen envisage a pairing of electrons in oxides that behave like metals and show superconductivity, much stronger than in the Cooper pairs proposed by Bardeen, Cooper and Schrieffer in their famous paper² of 1957 (the BCS theory). Many, but not all, use the concept of the bipolaron, known for the last decade³. In this article I shall first sketch the role of Cooper pairs in normal superconductors, then describe polarons and bipolarons, and finally I shall indicate how bipolarons and other forms of pairing can lead to superconduction at relatively high temperatures.

In a metal, or metallic oxide that behaves like a metal, the electrons are supposed to occupy quantum-mechanical states (but with no more than two per state) up to a limiting energy, called the Fermi energy, E_F . This has been known since the late 1920s, and there are many ways of demonstrating this experimentally. The first theories neglected the interaction of the electrons with phonons (the quanta of lattice vibrations), but this is now known to produce several effects. The first is to scatter the electrons, leading to the familiar electrical resistivity of ordinary conductors that increases with temperature. The second, as can be shown by an elementary application of perturbation theory, is to increase slightly the effective mass of electrons in states near E_F . The third is to introduce a weak attractive force between electrons. This would not in itself be strong enough to enable them to bind in pairs, were it not for the interaction of all the other electrons, which through a sophisticated analysis makes this possible. As it is, however, electrons with energies near E_F (not throughout the whole band) can form Cooper pairs, with a very weak binding energy of the order of $\hbar\omega\exp\{-1/N(E_F)V\}$. Here ω is the phonon frequency, $N(E_F)$ the density of electron states near E_F and V an electron–phonon coupling constant. The pairs are bosons — which means that the condition that only two can occupy the same state is no longer valid; at 0 K they must all condense into the same, lowest state. For rather involved reasons this condensate of pairs can move as a whole, giving rise to an electric current which is not scattered by either phonons or impurities, and persists indefinitely.

Many theories of the high-temperature

superconductors seem to envisage pair formation by a simpler, stronger mechanism than that of Cooper, that extends to all the mobile electrons of the material. As long ago as 1973, P. W. Anderson proposed that in certain crystalline structures, pairs might form in a rigid (phononless) lattice held together by the same kind of force that is familiar in the hydrogen molecule, that they are mobile, and that a condensed Bose gas would be the lowest state of the system. In Anderson's most recent work⁴ these play the role of the Cooper pairs, but are much more strongly bound. If he is right, phonons play no role and these materials should show no change in the transition temperature if an isotope in the oxide were changed (which would change the phonon frequency). Such an experiment will be difficult until single crystals are available, because the transition temperatures for the sintered powders used now are not sharp. Rumour, however, has it that in one material the effect has been looked for and not found.

Other theories use the concept of the bipolaron. The polaron is usually described as an electron (or a positive hole) in a semiconductor that distorts the lattice around it. At a temperature T such that $kT < (1/2)\hbar\omega$ it moves by a thermally activated hopping process with a mobility proportional to a Boltzmann term $\exp\{-W_H/kT\}$; W_H may typically be ~ 0.2 eV. But at low T the polaron behaves as a very heavy particle, with effective mass $\sim 2m_e \exp\{2W_H/\hbar\omega\}$ and no activation barrier to hopping. In one-dimensional problems a mass enhancement of this kind will always occur, but in three dimensions it occurs only if the electron energy-bands are narrow or if the phonon–electron interactions are strong. In two-dimensional systems the condition for polaron formation is less exacting than in three.

I suggested many years ago⁵ that, for a 'metal' in which the mobile electrons lie in a narrow electronic band, for instance a d -band, the mass enhancement normally treated by perturbation theory could be large, and extend throughout the whole band; one could then describe the system as a degenerate gas of polarons. One system where this is certainly the case is doped KTaO_3 , a material with a large band-gap but high static dielectric constant κ . Only if this is so is the attraction between electron and donor given by $e^2/\kappa r$, and so small enough to give metallic conduction for the low concentrations of donors observed (10^{17} cm^{-3}). For other reasons it appears to be the case for $\text{La}_{1-x}\text{Sr}_x\text{VO}_3$, a ceramic that shows metallic

properties investigated by Professor Hagenmuller's group at Bordeaux. The field of the randomly distributed La^{3+} and Sr^{2+} ions appears to produce Anderson localization in the d -band of vanadium, and this is unlikely unless the effective mass of electrons in that band is enhanced⁶.

The first clear evidence for the existence of bipolarons, that is pairs of polarons stuck together that move as such, was obtained for Ti_2O_3 by Lakkis *et al.*³ at Grenoble. In this material there are equal numbers of Ti^{3+} and Ti^{4+} ions. At low temperature these order (as do the Fe^{2+} and Fe^{3+} in magnetite) and the material is non-metallic, but non-magnetic; it must therefore be supposed that it is pairs of Ti^{3+} ions, bound by homo-polar forces, that take up ordered positions. At 130 K the pairs undergo a 'Verwey transition' and the material conducts with a small activation energy. It remains diamagnetic, showing that the carriers are bipolarons, with the electron spins anti-parallel, their magnetic moments cancelling.

A classical disordered array of bipolarons is thus possible. The questions arise, is a Bose condensation of bipolarons possible, and would it be superconducting? Alexandrov *et al.*⁷ considered the possibility in some detail in 1986, and concluded that the answer to both questions was yes but not at the high temperatures found. As pointed out particularly by Ray⁸, one needs a material with not too wide a d -band but neither too narrow. If it is too wide, bipolarons will not form at all; if it is too narrow, then the effective mass will be too great and the system will order as in Ti_2O_3 under the influence of the electrostatic repulsion between the charge carriers. Ray points out that the width of a d -band in an oxide depends on the degree of hybridization of the $3d$ with $2p$ orbitals of the oxygen ions. By good fortune this must be just right for La_2CuO_4 doped with Ca, Sr or Ba for compositions in which the material is metallic at room temperature and which show the newly discovered high- T_c superconducting transition. One can speculate on whether Ti_2O_3 might show the phenomenon under a very high pressure.

Several authors, then, envisage a degenerate gas of paired electrons, held together by forces stronger than the BCS phonon-based attraction. Those pairs may

move as bipolarons, the two electrons drawn together closer than the lattice spacing⁹, though this is uncertain. If this is so, the fact that for example La_2CuO_4 has a layer structure (see for example the recent article by Davies and Tilley in *Nature*¹⁰) may be relevant, because bipolarons are expected to form more easily in two dimensions than in three. There are those

who suppose, however, that the only new feature in these materials is that optical (high-energy) phonons rather than acoustical (low-energy) ones are involved, for which $\hbar\omega$ is much greater to give a higher transition temperature. □

Nevill Mott is Emeritus Professor of Physics, Cavendish Laboratory, Cambridge University, Cambridge CB3 0HE, UK.

Chemical physics

Cooling of molecular clusters

Tony Stace

THE isentropic ($\Delta S = 0$) expansion of a gas at high pressure through a small aperture into a vacuum can lead to the onset of condensation and the formation of molecular aggregates, or clusters as they are more commonly known. Cluster formation is a non-equilibrium process, and as yet it is not possible to make quantitative predictions about cluster growth and temperature from such experimental

Calculated cluster temperature and observed phase for various substances.

Substance	Cluster phase	Melting point (K)	Boiling point (K)	T_{cl} (K)
Cyclopropane	l	145.5	240.4	116
<i>n</i> -Butane	l	134.9	272.7	133
Neopentane	l	256.6	282.6	136
Benzene	l	278.7	353.3	179
Carbon tetrachloride	l	250.3	349.9	175
Argon	s	83.8	87.3	36
Sulphur hexafluoride	s	222.5	206.3*	129
Perfluorocyclobutane	s	233.0	267.2	155
Carbon tetrachloride	s	250.3	349.9	208

Phase (l, liquid; s, solid) determined by electron diffraction; melting and boiling points of bulk substances; and temperature T_{cl} of the cluster calculated using Klotz's formula (from ref. 6). Asterisk, sublimation temperature.

parameters as the aperture diameter and gas pressure. On page 222 of this issue, Klotz¹ presents a simple prescription which allows the temperature of a cluster to be estimated from the molar heat of evaporation of the material under consideration. An important feature of the rationale behind his formula is that the clusters be evaporating; only then can the kinetics of such a process be used to infer a temperature.

Almost any substance that can exert an appreciable vapour pressure in the region of the aperture can be made to cluster, either with itself or with other gaseous material. Hence, argon and benzene will not only form homogeneous Ar_n and $(\text{C}_6\text{H}_6)_n$ clusters quite readily but also, under the right conditions, they can give $\text{C}_6\text{H}_6 \cdot \text{Ar}_n$. Depending on the experimental conditions, n can range from 2 to about 1,000. Recent developments, most noticeably by Smalley and colleagues² at

Rice University, have made possible the generation of clusters from refractory materials, such as silicon, tungsten and copper, by laser vaporization of the solid surface followed by isentropic expansion.

The experimental requirement of high gas pressure before expansion means that when passing through the aperture, the atoms or molecules experience many collisions and finally emerge into the vacuum all having approximately the same velocity. Motion through the aperture is best described in terms of hydrodynamic or bulk flow, with much of the relative kinetic energy present before the expansion becoming associated with the emergent flow. The result is an increase in the average velocity of the beam molecules and a narrowing of the velocity distribution. A decrease in the width of a velocity distribution can be identified with a drop in temperature, and temperatures of about 5–10 K are quite commonplace in expansion experiments. Under these circumstances, the small attractive (van der Waals) forces between atoms and molecules become important and clustering or condensation can occur.

In atoms the only cooling is translational, but for molecules rotational and vibrational cooling also occurs. Because of the varying efficiencies of these processes, however, the final temperatures of the various degrees of freedom are quite often very different, with the order usually being translational < rotational < vibrational. As the clusters emerge from the aperture and travel through the vacuum, there is a rapid decrease in density which is accompanied by a decrease in collision frequency. A point is reached where the clusters experience no further cooling resulting from collisions, and any further decrease in temperature (internal energy) has to come from evaporation. It is at this stage of a cluster's history that the expression of Klotz¹ becomes both useful and relevant. Approximately 1 per cent or less of the expanded gas goes to form clusters, the remainder acts as a heat (and entropy) sink. Although very low temperatures are generated within the expansion, clustering is still an exothermic process.

1. Maddox, J. *Nature* **325**, 756 (1987).
2. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. *Phys. Rev.* **106**, 162–164 (1957).
3. Lakkis, S., Schlenker, C., Chakraverty, B. K., Buder, R. & Marezio, M. *Phys. Rev.* **B14** 1429–1440 (1976).
4. Anderson, P. W. *Science* **235**, 1196–1198 (1987).
5. Mott, N. F. *Phil. Mag.* **20**, 1–21 (1969).
6. Mott, N. F. & Davis, E. A. *Electronic Processes in Non-Crystalline Materials*, 2nd edn (Oxford University Press, 1979).
7. Alexandrov, A. S., Ranninger, J. & Robaszkiewicz, S. *Phys. Rev.* **B33**, 4526–4541 (1986).
8. Ray, D. K. *Phil. Mag. Lett.*, in the press.
9. Kamumura, H. *Jap. J. appl. Phys.* (in the press).
10. Davies, A. H. & Tilley, R. J. D. *Nature* **326**, 858–861 (1987).

The attempt to define a temperature for a cluster has associated with it a certain ambiguity. The experimental conditions are such that there may well be a distribution of temperatures. But, more important, temperature is an equilibrium concept maintained by collisions within (in the language of statistical mechanics) a canonical ensemble (constant volume, number of atoms and temperature). In contrast, once a cluster enters the rarified and isolated environment of the vacuum, it becomes a microcanonical ensemble (also constant volume and number of atoms, but constant total energy, not temperature). Klotz attempts to derive a temperature based on the kinetic component of a cluster's total energy, and evaporation (unimolecular decay) provides a measure of that kinetic energy. Clusters have no boundaries, and therefore, as Berry and co-workers³ noted recently, if they were treated as canonical ensembles then there would be nothing to stop the clusters from eventually losing all their atoms or molecules through evaporation.

Given the conceptual limitation discussed above, there is a use in the chemical physics of clusters for a measure of temperature. Cluster formation offers an opportunity to study at a microscopic level the intermolecular forces which are responsible for many of the properties of condensed matter. Experimental techniques, such as electron diffraction^{4,5}, have reached a stage where they can now distinguish between liquid-like and solid-like clusters. It therefore becomes important to know what physical state should be expected, or whether clusters can be generated in some other form, such as a supercooled liquid. The table shows some examples of temperatures estimated from the expression of Klotz, together with the physical state of the clusters as determined from an electron-diffraction experiment⁶. Those substances which in the bulk phase have narrow liquid-existence ranges (their melting and boiling points are close to each other) appear to give crystalline clusters. Their calculated temperatures are in keeping with that observation. For those examples where the cluster phase appears to be liquid-like, the temperature prediction suggests that they are supercooled. Of the molecular cluster systems studied so far, most appear to be in a liquid-like phase. There is evidence both from computer studies⁷ of argon clusters and from experiments on gold clusters⁸, that the freezing point of a finite-sized system of atoms can be significantly depressed. Whether this is the case for some of the molecular systems given in the table remains to be seen.

At present there is considerable research being conducted on clusters of all types. Particular attention is being paid to the possible catalytic properties of small

metal clusters. Apart from their high surface-to-volume ratio, it has been proposed that metal clusters of a certain size may provide surface sites which are particularly efficient at promoting catalytic activity. □

1. Klotz, C.E. *Nature* **327**, 222-223 (1987).
2. Dietz, T.G. *et al.* *J. chem. Phys.* **74**, 6511-6512 (1981).
3. Davis, H., Jellinek, J. & Berry, R.S. *J. chem. Phys.* (in the press).

4. Torchet, G., Schwartz, P., Farges, J., de Feraudy, M.F., & Raoult, B. *J. chem. Phys.* **79**, 6196-6202 (1983).
5. Valente, E.J. & Bartell, L.S. *J. chem. Phys.* **80**, 1451-1457 (1984).
6. Bartell, L. & Barshad, Y.Z. *J. phys. Chem.* (in the press).
7. Natanson, G., Amar, F. & Berry, R.S. *J. chem. Phys.* **78**, 399-408 (1983).
8. Buffat, Ph. & Borel, J-P. *Phys. Rev. A* **13**, 2287-2298 (1976).

Tony Stace is at the School of Chemistry and Molecular Sciences, Sussex University, Falmer, Brighton BN1 9QT, UK.

Palaeoanthropology

Who is the 'real' *Homo habilis*?

Bernard Wood

ON page 205 of this issue¹, Johanson and colleagues report the discovery of a 1.8-million-year-old skull and partial skeleton from Olduvai Gorge, Tanzania. More than 60 hominid specimens have now been recovered from this site, most coming from the lowermost deposits of Bed I and lower Bed II. These remains have been assigned to either *Homo habilis* or *Australopithecus*



The site at Olduvai Gorge where Johanson and colleagues found the hominid¹. White tags mark the location of bones found on the surface. The bones and fragments of the specimen are arranged in anatomical fashion on the floor of the excavation.

boisei. The former is the earliest known species of our own genus, the latter is a 'robust' australopithecine known from several East African sites. The new find rudely exposes how little we know about the early evolution of *Homo*.

Opinions about *H. habilis* differ. Some see it as a variable, but well-defined, taxon, whereas others have come to acknowledge that the label *H. habilis* has been applied to such a heterogeneous collection of material that the identity of the 'real' *H. habilis* is now all but obscured. Johanson and colleagues believe their specimen, OH62,

provides evidence that consolidates the taxonomic unity of the *H. habilis* hypodigm, or reference sample. An alternative interpretation is that OH62 confirms that the range of variation within material from the early Pleistocene of East Africa assigned to early *Homo* is now too great to be sensibly encompassed within one taxon.

The holotype, or defining specimen, of *H. habilis* is OH7, which was discovered at site FLKNN at Olduvai in 1960. It comprises part of the cranial vault, the damaged mandible and hand bones of a juvenile, and together with four other specimens (OH4, 6, 8 and 13) from Bed I and lower Bed II, it formed the original hypodigm of *H. habilis*². Since 1964, a dozen or so additional specimens from Olduvai have been attributed to *H. habilis*, including two crania (OH16 and 24), a mandible (OH37) and a juvenile upper dentition (OH39). Sites elsewhere in Africa have also contributed to the sample of *H. habilis*, including material from Swartkrans³, Sterkfontein⁴, the Omo⁵ and, most notably, Koobi Fora⁶. Despite the enlarged hypodigm, the main present-day doubts about *H. habilis* (does it represent more than one taxon and, if so, is the second taxon *Homo* or a 'gracile' australopithecine?) are the same as the ones that palaeoanthropologists debated in *Nature* more than 20 years ago. Robinson⁷ considered that the original *H. habilis* hypodigm was a conflation of *Australopithecus africanus* and *Homo erectus*, whereas Louis Leakey⁸ saw *H. habilis* as possibly embracing two species of *Homo* from separate lineages, one leading to *Homo sapiens* and the other to *H. erectus*. More recent contributions have suggested different splits of the *H. habilis* hypodigm (for example, ref. 9) but all these commentators are in agreement that, as it is presently understood, *H. habilis* is too variable to be a 'sound' species. To these difficulties have to be added differences over functional interpretations, in particular those of locomotor capability. Is the postcranial skeleton of *H. habilis* essentially human in both form and function¹⁰, or do the remains betray a less than modern human-like adaptation to bipedalism¹¹?

It is against the background sketched out above that putative *H. habilis* specimens have to be considered. The authors of the new paper¹, which assesses the affinities of OH62, put forward compelling evidence for it not being a 'robust' australopithecine. But how convincing is their evidence for it being placed in the genus *Homo*, and not in *Australopithecus*? Their case rests in large measure on the affinities between OH62 and Stw53, a skull recovered from Member 5 at Sterkfontein. The logical 'trail' becomes tenuous because Stw53 has merely been likened to *H. habilis*, and not formally attributed to it, even though more than a decade has elapsed since its discovery. Features of the dentition of Stw53 are decidedly primitive: 3-rooted upper premolars, and although its facial morphology is one of its most 'hominine' features, acknowledged experts have admitted that the facial skeleton of *A. africanus* (the most likely alternative taxonomic attribution for Stw53) is itself variable in form¹².

Even if we set aside the possibility that OH62 may belong to *Australopithecus*, and accept its allocation to *H. habilis*, does its inclusion in that species confirm or weaken the integrity of that taxon? Many attempts to sort within *H. habilis* have concentrated on absolute size, but some studies, especially those which have looked at variation within early *Homo* from Koobi Fora, have also looked at shape differences^{13,14}. Nonetheless, both approaches have suggested a broadly similar split of *H. habilis*, with one subgroup with large brains, orthognathic faces and relatively large dentition (such as KNM-ER1470 and 1590), and the other being smaller-brained with more lightly built faces (KNM-ER1813 and OH24). The elements of the face of OH62 that are preserved appear to resemble the small subgroup. It is natural to suspect that these two subgroups may be sexual dimorphs, but the pattern of variation within and between them does not suggest this. In summary, the cranial evidence of OH62 does little to unite the two morphs within *H. habilis*, nor does it particularly help to resolve whether they represent intraspecific, and presumably sexual, differences, or taxonomic variation.

Perhaps the most interesting aspect of OH62 is the light it sheds on the possible range of variation in the postcranial skeleton of early *Homo*. The shape and the size of the proximal femur, and the anatomy and relative lengths of the limb bones, both run counter to the view which sees *H. habilis* as a biped with a postcranial skeleton that is essentially modern human in its morphology, proportions and, by inference, function. Could the postcranial skeleton of OH62 be accommodated within the taxon represented by the femoral remains from Koobi Fora, such as KNM-ER1472 and 1481, when the latter are apparently sufficiently modern human to have been attributed by some to *H. erectus*?

The answers to these and the many other intriguing questions raised by OH62 will come only when field workers have provided sufficiently large samples, and when the deskbound analysts have refined their knowledge of early hominid variation. □

1. Johanson, D.C. et al. *Nature* **327**, 208–209 (1987).
2. Leakey, L.S.B. et al. *Nature* **202**, 7–9 (1964).
3. Clarke, R.J. & Howell, F.C. *Am. J. phys. Anthropol.* **37**, 319–336 (1972).
4. Hughes, A.R. & Tobias, P.V. *Nature* **265**, 310–312 (1977).
5. Boaz, N.T. & Howell, F.C. *Am. J. phys. Anthropol.* **46**, 93–108 (1977).
6. Howell, F.C. in *Evolution of African Mammals* (eds Maglio, V.J. & Cooke, H.B.S.) 154–248 (Harvard

- University Press, Cambridge, Massachusetts 1978).
7. Robinson, J.T. *Nature* **205**, 121–124 (1965).
8. Leakey, L.S.B. *Nature* **209**, 1279–1281 (1966).
9. Leakey, R.E.F. & Walker, A. *Science* **207**, 1103 (1980).
10. Day, M.H. In *Hominid Evolution. Past, Present and Future* (ed. Tobias, P.V.) 115–127 (Liss, New York, 1985).
11. Lewis, O.J. *J. Anat.* **131**, 275–298 (1980).
12. Rak, Y. *The Australopithecine Face* (Academic, New York, 1983).
13. Wood, B.A. in *Ancestors: The Hard Evidence* (ed. Delson, E.) 206–214 (Liss, New York, 1985).
14. Stringer, C. in *Major Topics in Primate and Human Evolution* (eds Wood, B., Martin, L. & Andrews, P.) 266–294 (Cambridge University Press, 1986).

Bernard Wood is Derby Professor of Anatomy at Liverpool University, PO Box 147, Liverpool L69 3BX, UK.

Cell signalling

G proteins and diabetes

Stephen R. Pennington

SOME of the known guanine nucleotide-binding proteins, or G proteins are part of the machinery that provides a link between the outside of a cell and its interior. Attention has recently been focused on these proteins both because of their critical involvement in cell signalling processes and because of the continued discovery of novel G proteins (see refs 1, 2). On page 229 of this issue³, Gawler, Houslay and colleagues describe the first evidence implicating the dysfunction of a G protein in a pathological state.

The best understood of the G proteins couple the enzyme adenylate cyclase to a wide variety of cell surface hormone receptors. Adenylate cyclase produces cyclic AMP which initiates a complex amplification cascade of phosphorylation reactions to generate the intracellular response characteristic of the hormone. Of the two G proteins that interact with adenylate cyclase, one mediates a stimulatory input (G_s) and the other the inhibitory input (G_i)⁴. In turn both G_s and G_i are composed of three subunits, α , β and γ . The β and γ subunits of G_s and G_i are very similar whereas their α -subunits are unique². Houslay and co-workers demonstrate³ that in a laboratory animal model of type I diabetes the inhibitory activity of G_i on adenylate cyclase is lost. The pivotal role of this G protein in normal cell signalling suggests that the observed loss of G_i

accounts for at least some of the dramatic consequences of the disease.

The mechanism by which a trimeric G protein activates its target effector system is shown in Fig. 1. The extracellular hormone binds to its receptor, allowing the association of the receptor molecule with a G protein (Fig. 1b), which results in the replacement of non-covalently bound GDP in the α -subunit with GTP which in turn promotes the dissociation of the G protein. The released α -subunit, with GTP bound, can interact with the effector system (adenylate cyclase, for example) and activate or inhibit it (Fig. 1c). In time the GTP bound to the α -subunit is hydrolysed to GDP and, in this form, the α -subunit can reassociate with the $\beta\gamma$ -dimer, so regenerating the inactive G protein (Fig. 1a). The α -subunits therefore seemingly play the most important role in G-protein function. The $\beta\gamma$ -subunits may anchor the α -subunit in the membrane (see ref. 2). Additionally, the common $\beta\gamma$ -subunits may enable 'cross-talk' between the G proteins⁴.

Houslay and colleagues³ use two independent measures of G_i function to demonstrate that when rats are made diabetic by drug treatment their livers show a decreased inhibitory input to adenylate cyclase. The inhibitory activity of G_i is restored by the administration of insulin to the diabetic animals, suggesting

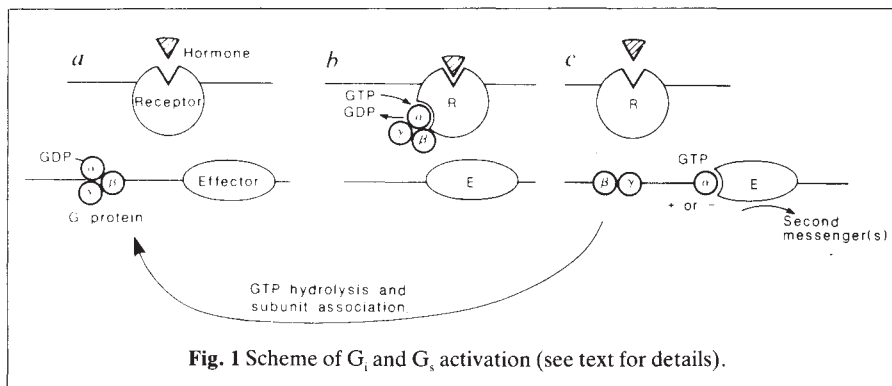


Fig. 1 Scheme of G_i and G_s activation (see text for details).

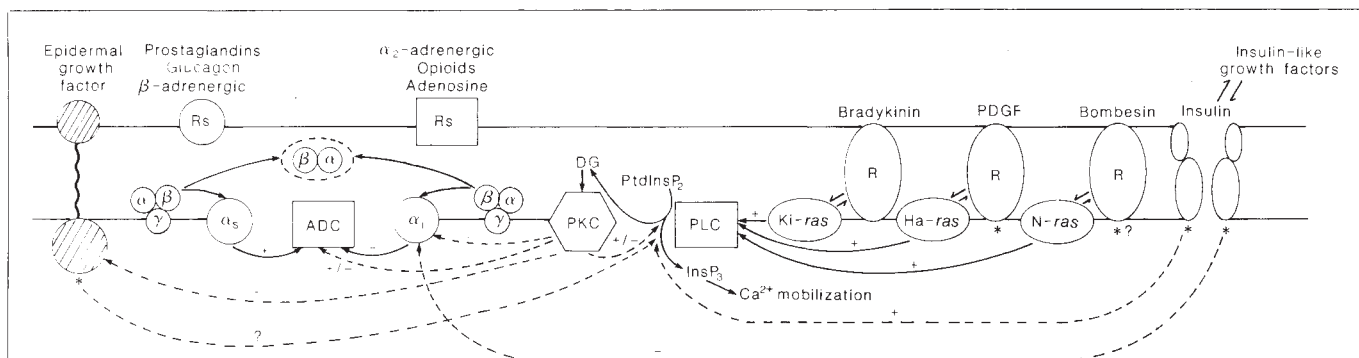


Fig. 2 Possible sites of interaction between signalling pathways (broken lines). Interaction is at several levels: (1) within G proteins by common $\beta\gamma$ -subunits; (2) phosphorylation of adenylate cyclase, G_i and/or receptors by protein kinase C; (3) phosphorylation of G and/or G γ by receptor kinases; and (4) between receptors, that is, cross-reactivity between insulin and the insulin-like growth factors. PLC, phosphatidyl bisphosphate phospholipase C; InsP_3 , inositol trisphosphate; DG, diacylglycerol; PKC, protein kinase C; PtdInsP_2 , phosphatidylinositol 4,5-bisphosphate; R, receptor; PDGF, platelet-derived growth factor; ADC, adenylate cyclase; *, receptor tyrosine kinase.

that the concentration of insulin circulating in the blood (which is low in the experimentally induced diabetic animals) controls the activity of G_i in the liver. The authors go on to show that the loss of G_i activity is associated with a reduction in the amount of the α -subunit of G_i in liver membranes measured by immunochemical techniques. As anticipated, insulin therapy can partially reverse the depletion of G_i α -subunit. This latter observation highlights the value of antibody-detection techniques in identifying the different G proteins. Two bacterial toxins (cholera toxin and pertussis toxin) are often used to identify and modify the function of G proteins, but it has become increasingly clear that these toxins each interact with several G proteins, so making it difficult to define specifically an observed response to a particular G protein.

How might insulin regulate the level of the α -subunit of G_i ? There are many possibilities; one attractive mechanism might involve the direct phosphorylation of this subunit by the insulin receptor tyrosine kinase⁵ (Fig. 2) with the resultant removal of the modified protein from the membrane.

What of the other G proteins? Recent evidence implicates the *ras* family of G proteins in the coupling of hormone receptors to phosphatidylinositol bisphosphate hydrolysis⁶ (Fig. 2). This reaction results in the generation of two putative second messengers⁷, one, inositol trisphosphate, which causes an increase in the Ca^{2+} concentration within the cell, the other (diacylglycerol) activates protein kinase C. The activation of this kinase, as well as resulting in the phosphorylation of many cellular enzymes, may also initiate another level of crosstalk. Thus, protein kinase C can phosphorylate and down-modulate receptor binding at the epidermal growth factor receptor, and it has also been shown that protein kinase C can directly phosphorylate the catalytic subunit of adenylate cyclase⁸ as well as the α -subunit of G_i ⁹.

Immunological estimates of the total G-protein content of cells show that there may still be many unidentified G proteins¹⁰. Identifying these proteins, as well as deciphering the details of the complex interactions between the different membrane signalling processes, is the challenge for the future. The most valuable reward will be the understanding of the changes in these signalling events that occur in diseased states. □

1. Bourne, H.R. *Nature* **321**, 814–816 (1986).
2. Bourne, H.R. *et al. Biochem. Soc. Trans.* **15**, 35–38 (1987).
3. Gawler, D., Milligan, G., Spiegel, A. M., Unson, C. G. & Houslay, M. D. *Nature* **327**, 229–232 (1987).
4. Levitski, A. *FEBS Lett.* **211**, 113–118 (1987).
5. O'Brien, R. M. *et al. FEBS Lett.* **212**, 281–288 (1987).
6. Berridge, M.J. & Irvine, R.F. *Nature* **312**, 315–321 (1984).
7. Wakelam, M. J. *et al. Nature* **323**, 173–175 (1986).
8. Yoshimasa, T. *et al. Nature* **327**, 67–70 (1987).
9. Katada, T. *et al. Eur. J. Biochem.* **151**, 431–437 (1985).
10. Milligan, L. G. *et al. FEBS Lett.* **195**, 225–230 (1986).

Stephen R. Pennington is in the Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK.

Immunoglobulin genes

Circular T-cell receptor gene recombination products

Frederick W. Alt and George D. Yancopoulos

ON page 242 of this issue¹, Fujimoto and Yamagishi describe the use of a novel circular DNA isolation technique to isolate and characterize excision products of T-cell receptor (TCR) α -chain gene variable (V) region recombination events in the normal mouse thymus. In a separate report², another group has similarly analysed the excision products from joins involving TCR β -chain V-region gene elements. Several years ago, Yamagishi and his colleagues noted a qualitative difference in extrachromosomal DNA circles isolated from primary lymphoid organs relative to other tissues and speculated that such circles in the thymus derive from rearrangements of the then uncharacterized TCR genes³. The results of the new studies confirm that speculation and extend ideas about the mechanism by which immunoglobulin or TCR V-region gene segments are assembled.

Immunoglobulin and TCR V-region genes are assembled from component gene segments that are separated in the germ line and joined during the differentiation of the respective B and T-lymphocyte lineages. There are three families of immunoglobulin genes (heavy

chain; κ - and λ -light chains) and at least three families of TCR genes (α , β and γ). Each of these families has a distinct chromosomal location and the assembled V-region genes consist either of two or three joined segments: V, diversity (D) and joining (J) segments; or just V and J segments. The joining of these different gene segments seems to be targeted by conserved nucleotide sequences (recombination recognition sequences) that flank the side of each germline segment; these recognition sequences consist of a palindromic heptamer (open triangles in the figure) and a characteristic nonamer (closed triangles) separated by a non-conserved 'spacer'. These conserved recognition sequences flank all known immunoglobulin or TCR V-gene segments and are therefore probably the target of a common site-specific recombination system.

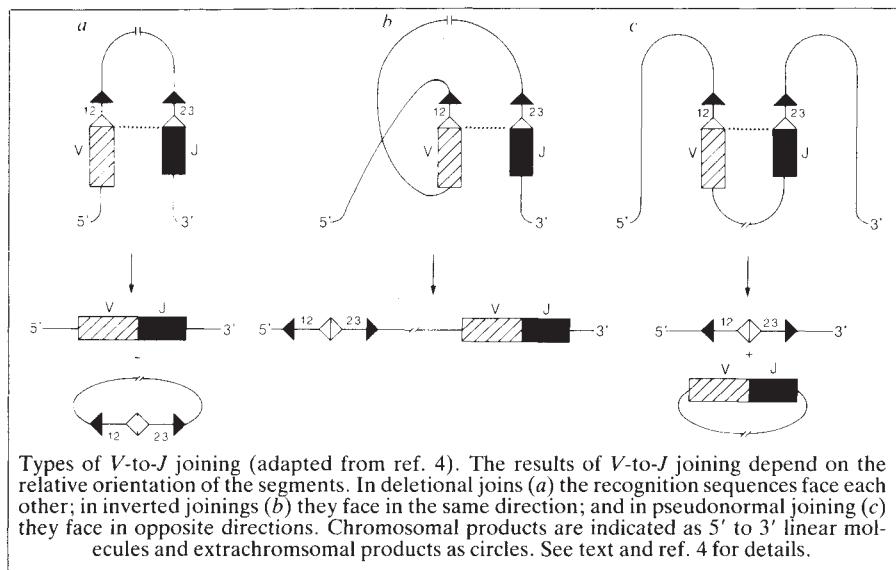
Based on analyses of an unusual immunoglobulin heavy-chain D–J joining event and the products of other types of joins (see below), a general model for this site-specific recombination was proposed⁴. According to this model, joining occurs by a two-step non-reciprocal recombination in which the first event is a

precise double-stranded break between the elements to be joined and their flanking recognition heptamers, followed by the precise (back-to-back) ligation of the two heptamers. Then, in a second step, the coding elements are joined in an imprecise event in which bases are often lost from one coding partner or both (*V* and/or *J*) and bases can sometimes be added *de novo* (without a template) between the two coding segments (*N* regions). Alternatives to this basic model have been outlined that differ in some details⁵, but that agree with the conclusion that the heptamers are generally fused precisely and the coding partners are not.

joining leads to inversion or deletion is simply based on the relative orientation of the participating segments in the chromosome (*a* and *b* in the figure). Therefore, it is clear that there is no *a priori* need to consider deletional or inversional joining mechanistically different. In fact, according to this model the apparently gratuitous circular deletion of fused coding segments with retention of the fused heptamers was predicted in the case of joining between elements whose recognition sequences were facing away from each other (so-called pseudonormal joining; ref. 4 and *c* in the figure). Significantly, the predicted products of such

(immunoglobulin κ -chain and TCR β -chain) use inverted joinings in normal *V*-gene assembly. In these cases, non-fusion of heptamers would result in a disastrous chromosomal fragmentation. Thus, the deletion of circles could be a byproduct of a recombination mechanism that has evolved to allow for both deletional and inversional joining events. It is also possible that circles are less recombinogenic with respect to chromosomal integration than linear molecules and thus minimize potential mutagenic damage.

The extrachromosomal, circular DNA isolation technology can be used to address many important and exciting questions. If the frequency of circles can be related to endogenous recombination rates, then this system could be used to quantitate *in vivo* the recombination system at various loci, in various lineages and at various developmental stages. Mechanistically, the system could be used to address questions not easily approached with other techniques. For example, although apparently site-specific *V*-gene replacement events have been described in cell lines^{10,11}, the relevance of such joining *in vivo* remains unknown because the chromosomal products of such joins are difficult to distinguish from normal *V-D-J* rearrangements. The predicted extrachromosomal reciprocal products of such joins should be readily identifiable and their relative frequency presumably can now be assayed. Another interesting mechanistic question derives from the observation that differentiating lymphocytes in mice homozygous for the *SCID* mutation apparently make abortive attempts to form *V*-region joins¹². Analysis of the extrachromosomal products of such abortive joins could elucidate the potential mechanistic consequences of the *SCID* defect with respect to the recombination system. Finally, and perhaps of greatest interest, the high frequency with which the TCR α -gene products occur in the preparation of Fujimoto and Yamagishi¹ strongly implies that this system can be used to identify novel sets of rearranging genes in lymphocytes and potentially even in other lineages. □



In contrast to previously analysed fused heptamers and those in the circles isolated by Yamagishi and colleagues (ref. 1 and personal communication), many of the circles characterized by Sakano and colleagues⁷ have several extra bases between the fused heptamers. These findings might require some modification of the existing mechanistic ideas, but could also result from a small or biased sample size.

That deletion of intervening DNA accompanies joining at various immunoglobulin or TCR loci has been extensively documented (see refs 6,7) and has even been exploited to map the relative order of segments along the chromosome. The occurrence of inverted joining between endogenous chromosomal *V*-gene segments was first demonstrated by the analysis of an inverted heavy-chain *D*-to-*J* join¹ made possible by the extensive characterization and linkage of the various segments in that region of the chromosome⁸. Inverted joining of a TCR β -chain *V* and *D*-*J* segment has been demonstrated⁹ and inverted κ light-chain *V*-*J* joining strongly implied⁷. Various references have been made to deletional or inversional models of *V*-gene assembly that imply different (distinct) mechanisms. Given the recombination model we have outlined above, whether or not

pseudonormal joins were identified in circular thymus DNA².

The occurrence of perfectly fused 'reciprocal' recognition heptamers as a result of the known or presumptive inverted joins outlined above was the impetus for the suggestion that excision products of deletions would be circular because of a similar heptamer fusion⁴. The new studies^{1,2} provide the first direct proof that the intervening DNA can be excised as a circle resulting from heptamer fusion during deletional joining; moreover, the work shows that these joins occur at a discernable frequency in the animal. But note that if linear DNA sequences were deleted at a significant or even greater frequency, they would not be detected by this system, which assays only for circles.

Assuming that heptamer fusion resulting in circles is the prevailing mechanism during deletional joining, then one must try to understand why this mechanism evolved. One possibility is that initial fusion of the heptamers evolved to ensure that the proper strands are joined during the subsequent complicated joining of the coding strands, thus avoiding potential chromosomal breaks or other aberrations. This consideration takes on further significance when one considers that at least one and probably two different loci

1. Fujimoto, S. & Yamagishi, H. *Nature* **327**, 242-243 (1987).
2. Okazaki, K., Davis, D. & Sakano, H. *Cell* (in the press).
3. Tsuda, T. et al. *Plasmid* **10**, 235-241 (1983).
4. Alt, F. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
5. Lewis, S., Gifford, A. & Baltimore, D. *Science* **228**, 677-685 (1985).
6. Sakano, H. et al. *Nature* **280**, 388-394 (1979).
7. Cory, S., Adams, J. & Kemp, D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4943-4947 (1980).
8. Kurosawa, Y. & Tonegawa, S. *J. exp. Med.* **155**, 201-208 (1981).
9. Mallisen, M. et al. *Nature* **319**, 28-33 (1986).
10. Reth, M. et al. *Nature* **322**, 840-842 (1986).
11. Kleinfeld, R. et al. *Nature* **322**, 843-846 (1986).
12. Schuler, W. et al. *Cell* **46**, 963-972 (1986).

Frederick W. Alt and George D. Yancopoulos are in the Departments of Biochemistry and Microbiology, College of Physicians and Surgeons of Columbia University, 630 W 168th Street, New York, New York 10032, USA.

Plate tectonics

When continents rift

Robert S. White

THE Tertiary volcanic region in the north-east Atlantic is a classic area for detailed petrological studies of magmatic processes. It is also important for detailed studies of large-scale geodynamic processes that lead to the generation of new ocean basins. When Greenland rifted away from Europe to form the north-east Atlantic, there was an immense outburst of volcanic activity both on the rifted continental margins and in the hinterland, creating the Tertiary igneous province of Greenland, the Faeroes, north-west Scotland and Ireland. The full extent, timing and magnitude of the offshore volcanism that occurred became apparent at a recent conference*.

The north-east Atlantic is a superb natural laboratory in which to study the birth of an ocean. The rifted margins mostly have little sediment cover, so that their deep structure is not obscured. Both sides of the rifted basin can be studied in a way not possible on many other margins and there is good control on the dating of major events. In the early 1980s it was realized that there are extensive volcanic flows on some rifted margins. Although the volcanic flows are offshore and therefore inaccessible to the geological hammer, they can be readily identified on seismic reflection profiles. They produce a series of layered reflections that dip seawards (Hinz, K. *Geol. J.* E22, 3-28; 1981). Dipping reflectors have now been recognized around the entire north-east Atlantic (see figure) and are known from drilling to be formed of basalt extruded near or above sea level. It is widely accepted by analogy with similar stacks of dipping lavas now forming in Iceland (Palmason, G. *J. geophys. Res.* 47, 7-18; 1980) that the dipping-reflector sequences were created by basalt flows from fissures to seaward of the margin (J. Mutter, Columbia University, USA).

There is considerably less agreement on the location of the continent-ocean boundary. According to Mutter, the dipping reflectors are formed entirely by sea-floor-spreading processes with the continental boundary to landward. Others think that they overlie stretched, intruded continental-crust (D. Roberts, British

Petroleum, London), a view supported by isotope evidence of continental contamination of the lavas (L. Parson, Institute of Oceanographic Sciences, UK; R. Merriam, British Geological Survey). Research by my group, along with colleagues at Birmingham and Durham Universities, shows that in the Rockall area often an

increasingly constrained by many new studies, including biostratigraphy around ash layers in the North Sea (R. Noble, British Geological Survey), radiometric dating (A. Mussett, University of Liverpool; I. Meighan, Queen's University, Belfast), palaeomagnetic measurements, magnetic-anomaly identification and seismic-reflection stratigraphy. The entire margin split very rapidly, with the bulk of the magmatism occurring in 2-3 million years or less.

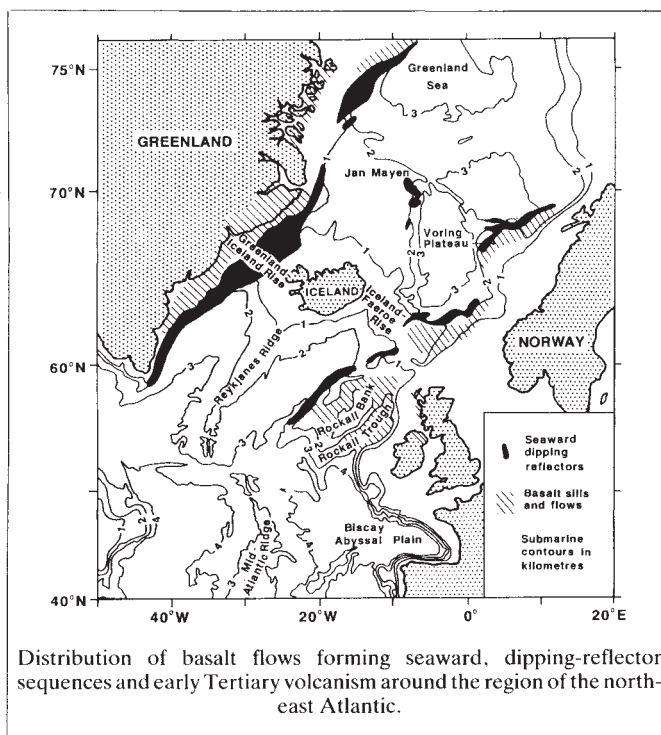
Huge volumes of igneous rock were generated during rifting. According to Roberts, the dipping reflectors alone account for $1-2 \times 10^6 \text{ km}^3$ of basalts; this is about the same as the entire Deccan volcanic province in India. Furthermore, new seismic-refraction data of Mutter and co-workers from the Voring Plateau, and our own from the Rockall margin show that up to five times more igneous rock is emplaced beneath the dipping reflectors, giving a total volume of $5-10 \times 10^6 \text{ km}^3$.

Why should there have been such intense and rapid igneous activity along this particular 2,000-km-long segment of continental rift? Mutter proposes small-scale convection under the evolving rift as the mechanism for producing thick igneous crust on some rifted margins. My view is that the large quantities of igneous rock on the rifted margin are due to anomalously hot material in the ductile asthenosphere that lies roughly below 100 km. As the rigid lithosphere above is stretched and

thinned in the initial stages of rifting, the hot asthenosphere wells up to fill the space, generating partial melt as it rises and decompresses. D. McKenzie and M. Bickle (University of Cambridge) have shown that 15-20 km of melt can be produced by upwelling asthenosphere that is only 100-150°C hotter than normal. These sorts of anomalous temperature in the asthenosphere are created across 1,500-2,000-km-wide regions by hot spots such as the one at present centred under Iceland. M. Bott (University of Durham) introduced a global perspective by postulating that the Icelandic hot spot is driven from the lower mantle by the return flow of material subducted into the lower mantle at the edge of the Pacific plate.

It appears that the entire Tertiary igneous province and elevated, volcanically active, continental margins of the north-east Atlantic can be explained by a relatively small ($\sim 100^\circ\text{C}$) increase in the asthenosphere temperature. □

Robert S. White is in the Department of Earth Sciences at the Bullard Laboratories, University of Cambridge, Cambridge CB3 0EZ, UK.



Distribution of basalt flows forming seaward, dipping-reflector sequences and early Tertiary volcanism around the region of the north-east Atlantic.

upper set of the reflectors overlies stretched continental crust but a seaward set occurs in the oceanic crust. Numerous schemes for defining the location of the continent-ocean boundary have been proposed: midway beneath the innermost dipping wedge (J. Skogseid, University of Oslo); at the seaward margin of the innermost wedge (D. Smythe, British Geological Survey) at the landward margin of the innermost wedge (Mutter); and where the lowest of the dipping reflectors stops onlapping (H. Larsen, Greenland Geological Survey).

Much of this debate on the exact location of the continent-ocean boundary is sterile because it revolves around poorly defined terminology. Between the unstretched, continental crust and the new, oceanic crust there must be a region where the percentage of igneous rock in the crust increases due to both intrusion and extrusion. It is meaningless to try to identify a precise boundary in this intermediate region of fragmented continental-blocks and newly emplaced igneous rock.

The timing of the volcanism is becoming

*Early Tertiary Volcanism and the Opening of the North-East Atlantic. Joint meeting of the Geological Society of London with the Norsk Petroleumsforening, 18-19 March 1987.

Climatology

Tropical Pacific variations

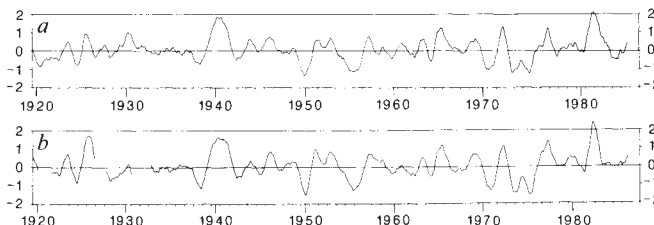
Eugene M. Rasmusson

THE tropical Pacific has been the focus of many recent studies of year-to-year climate variability associated with the El Niño–Southern Oscillation, or ENSO, phenomenon^{1–3}. The more difficult but equally important task of documenting lower-frequency variability has received far less attention. The potential for man-made climate change adds urgency to the task of establishing the characteristics of the natural climate noise background from which the man-made signals will emerge. On page 216 of this issue⁴, Kim Whysall and colleagues describe the results of a study of low-frequency variability in the tropical Pacific, based on a careful processing and analysis of millions of merchant-ship wind observations since 1920.

Until recently, most of the information on decade-scale climate variability was derived from land-based observations. Surface marine data acquired during this century have now been assembled into massive datasets suitable for climate analyses, such as the UK Meteorological Office Main Marine Databank used by Whysall and colleagues in their study. Unfortunately, changing instrumentation, observational procedures and techniques, together with major changes in the distribution of the observations, result in a far from homogeneous dataset. So questions concerning the reliability of observed trends must be addressed in evaluating analysis results, and this has hindered studies of low-frequency variability. Recent studies of climate trends^{5,6}, however,

have made extensive use of the sea-surface temperature observations.

There has been a gradual shift in surface wind measurements, from visual estimates based on the state of the sea (the Beaufort scale) to data obtained from anemometers, now mounted as much as 30–40 m above the ocean surface (V. Cardone, personal communication).



Time series of *a*, surface pressure anomalies for Darwin (12.4°N, 130.9°E); and *b*, the pressure difference between Darwin and Tahiti (17.5°S, 149.6°W). Departures, shown on the ordinate, are normalized by dividing the individual anomaly by the standard deviation of all anomalies for the individual calendar month. The pressure difference is again normalized in the same manner. Values are then smoothed with a 12-month running mean.

Aware of the uncertainty this causes, Whysall and colleagues have organized and summarized the historical wind data in ways that provide a clear documentation of the multi-year variability in the trade-wind systems, as revealed by the marine dataset. The analyses reveal a coherent pattern of variability on the scale of the Pacific basin.

The authors make a strong argument that the low-frequency variations revealed by the data are real. But to establish firmly the credibility of these or any other results derived from the imperfect climate data, it is important to examine their consistency with independent information: surface pressure, sea level and surface temperature data. The authors point out consistent relationships between the analysed wind variations and simultaneous variations in surface temperature associated with ENSO episodes. It is also useful to compare their wind-data time series with two widely used indices of Southern Oscillation variability — the surface pressure at Darwin, Australia, and the pressure difference between Darwin and Tahiti. These time series, shown in the figure, reflect year-to-year pressure fluctuations of opposite sign between the south-east Pacific and the west Pacific–Indian Ocean monsoon region, which are associated with the Southern Oscillation. The pressure variations at these points are an index of the changing equatorial pressure gradient associated with the changes in strength of the equatorial Pacific easterlies shown in Fig. 1 of Whysall *et al.*

Comparison of the pressure indices in the above figure with their Fig. 1, and with the trade-wind time series shown in their Fig. 2, reveals some interesting points.

First, all the series reflect the collapse of the tropical Pacific easterlies around 1940. Southern-Oscillation indices, however, show a complex ENSO event during 1939–41, and a separate weaker event in 1944, whereas the trade-wind series (their Fig. 2) show a more sustained, off-equatorial diminution of the easterlies that seems to last for five years or more.

Second, contrary to the North Pacific trade-wind series (their Fig. 2*a*), the Southern-Oscillation pressure indices do not imply a pronounced increase in the equatorial easterlies during the late 1940s, as compared with their strength during the years preceding the climatic event of the early 1940s. Neither do the indices show the marked trend towards increasing easterlies that is particularly pronounced in the South Pacific trade-wind data between 1950–80, as illustrated by their Figs 2 and 4. This apparent contradiction could result from the fact that the two pressure-index stations are west of the primary

area of diminished trade winds, in a region where equatorial wind data are sparse, and trends in the off-equatorial winds appear to have been more meridional (north–south) than zonal. But consistency cannot be ensured without additional analyses.

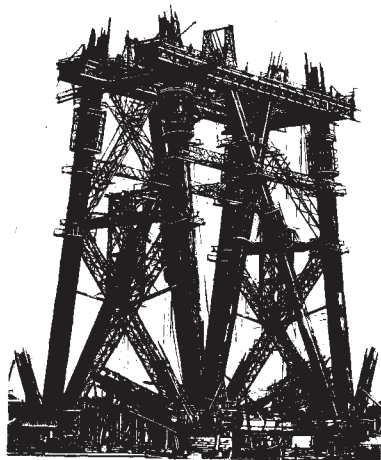
The difference and similarities between these time series illustrate the difficulty of establishing the pattern of decade-scale climate variability from the limited climate datasets now available and highlights the need to integrate as much information as possible in these analyses. The connection between year-to-year ENSO fluctuations and lower frequency variability is obscure. ENSO evolves primarily from coupled ocean–atmosphere interactions in the tropical Pacific, but extratropical ocean variability, as well as other global forcing mechanisms, are likely to be fundamental in lower-frequency variability, as noted by the authors. The paper by Whysall and colleagues provides benchmark analyses which should stimulate further analyses and exploitation of the rich source of climate information represented by the surface marine dataset. □

1. Rasmusson, E.M. *Am. Sc.* 73, 168–177 (1985).
2. Gill, A.E. & Rasmusson, E.M. *Nature* 306, 229–234 (1983).
3. Philander, S.G.H. *Nature* 302, 295–301 (1983).
4. Whysall, K.D.B., Cooper, N.S. & Bigg, G.R. *Nature* 327, 216–219 (1987).
5. Folland, C.K., Palmer, T.N. & Parker, D.E. *Nature* 320, 602–607 (1986).
6. Jones, P.D., Wigley, T.M.L. & Wright, P.B. *Nature* 322, 430–434 (1986).

Eugene M. Rasmusson is at the Department of Meteorology, University of Maryland, College Park, Maryland 20742, USA.

100 years ago

BRIDGING THE FIRTH OF FORTH



From *Nature* 36, 79; 26 May 1887.

Heterosexual transmission of AIDS by male drug users

SIR—May and Anderson state that “In developed countries at present and into the near future, it is probable that sexually-transmitted human immunodeficiency virus (HIV) infections among females are likely to come mainly from bisexual males”¹. Current epidemiological information from the United States, however, suggests that a substantial percentage of sexually-transmitted HIV infections among females will be derived from male intravenous drug abusers. It is important to appreciate the extent and transmission of HIV infection in the population of intravenous drug abusers and their sexual partners as this group accounts for a large proportion of the heterosexually-transmitted HIV infections and perinatal HIV transmissions^{2,3}.

Three parameters should be compared to assess the relative contributions of bisexual men and male intravenous drug abusers to the number of women who acquire HIV infection through sexual transmission.

First is the total pool of each. In the US state of Maryland for the year of 1986, 16.7% of AIDS (acquired immune deficiency syndrome) cases were bisexual men while 11.0% were male intravenous drug abusers; only 4.2% of the men with AIDS were both bisexual/homosexual and intravenous drug abusers⁴. Of all reported AIDS cases in the United States from before 1982 to 1986, 13.4% occurred in male intravenous drug abusers⁵ and although homosexuals are not distinguished from bisexuals it is likely that as in Maryland around 16.0% of all AIDS cases in the United States are bisexual men.

The second parameter is the number of new female partners each bisexual man encounters compared with the number of female partners each intravenous drug user encounters. Few data are available but in a survey of 212 heterosexual men without a specific history of intravenous drug abuse in San Francisco, 47.2% had two or more sexual partners in a six-month period of time, 12.2% had over four and 5% had none, whereas of 173 bisexual men in this survey only 32% had any female sexual partners during the same six-month period and 3.5% had over four⁶.

The third parameter is the rate of acquisition of HIV infection by sexual contact, which one could assume would be equal for both groups if one considers only vaginal intercourse and both groups were to use condoms to the same extent. There are insufficient data available to say whether bisexual men practice rectal intercourse with women more frequently than male intravenous drug abusers or whether rectal intercourse with women results in a higher rate of transmission

than vaginal intercourse⁶. Receptive rectal intercourse is a significant risk factor for HIV infection in homosexual men^{7,8}.

In summary, in the state of Maryland if only a third of bisexual men have sexual encounters with women during a specific period of time, the number of heterosexual drug abusers who may have sexual intercourse with females is greater. In addition, a larger percentage of male heterosexual intravenous drug abusers are likely to have a greater number of new female sexual partners than bisexual men. Therefore, it is probable that the number of sexually-transmitted HIV infections among females in the United States from male intravenous drug abusers will be at least as great as that from bisexual men.

PRISCILLA A. FURTH

University of Maryland Medical School,
Baltimore, Maryland 21201, USA

1. May, R.M. & Anderson, R.M. *Nature* **326**, 137-142 (1987).
2. Centers for Disease Control *MMWR* **34**, 721-726; 731-732 (1985).
3. Centers for Disease Control *MMWR* **35**, 757-760; 765-766 (1986).
4. *Maryland AIDS Update* 1-3 (Maryland State Department of Health and Mental Hygiene, 31 December 1986).
5. Winkelstein, W. Jr, Samuel, M., Padian, N.S. & Wiley, J.A. *J. Am. med. Ass.* **257**, 1470-1471 (1987).
6. Fischl, M.A. *et al. J. Am. med. Ass.* **257**, 640-644 (1987).
7. Mayer, K.H. *et al. Am. J. med.* **80**, 357-363 (1986).
8. Winkelstein, W. Jr *et al. J. Am. med. Ass.* **247**, 321-325 (1986).

AIDS vaccine strategies

SIR—It is universally agreed that development of a vaccine represents the most effective means to contain and prevent AIDS. Based on wide experience with vaccines against some acute viral infections, the scientific premise upon which that strategy seems currently to rest is that it will be possible to elicit antibodies that neutralize the human immunodeficiency virus (HIV) or otherwise prevent infection. The principal concerns that have been aired relate to the heterogeneity of the virus, the problem of identifying antigens and epitopes required for neutralization, the formidable problems associated with testing and evaluating candidate vaccines, and the vexing issues of liability in the developed world and availability in the developing world.

While everyone hopes that this research will lead to an effective vaccine, I believe it worth mentioning the concern of some immunologists that the basic premise may not be entirely correct. The observations that the genome of HIV is detectable in only 1 in 50,000–100,000 lymphocytes of patients with AIDS and that only low levels of virus or reverse transcriptase can be found in seminal fluid and serum, raise the possibility that infection is acquired not only, or even primarily, from free virus, but through transmission of already infected lymphocytes, macrophages or spermatozoa. The apparent ability of HIV-encoded proteins to induce cell

fusion and syncytia formation, particularly if an allograft reaction is mounted by the host against the intruding virus-infected cells, could enhance cell-to-cell transmission. This could greatly complicate the development of an effective vaccine because antibodies are generally ineffective in eliminating virus-infected or transformed cells. The immunologic problem would become how to eliminate already infected cells rather than to neutralize free virus. Currently the cell-mediated immune mechanisms involved in killing cells infected by acute and persisting viruses are less well understood than antibody-mediated phenomena, and they bring the complexity that the epitopes presented to T cells are at least as much a function of the individual's major histocompatibility antigens as of the structure of the viral proteins, and will vary from individual to individual.

In my judgement, recognition of this possibility would call for great caution in promising the public an effective vaccine in the short term, for intense efforts to ascertain whether the virus is transmitted primarily by infected cells or by free virus, and for greater consideration of cell-mediated immune strategies for developing immunity against virus infected cells. The hopeful aspect of these considerations is that if a vaccine can be developed that effectively and safely eliminates virus-infected cells, it could be effective therapeutically as well as prophylactically.

BARRY R. BLOOM

Department of Microbiology and
Immunology,
Albert Einstein College of Medicine,
Bronx, New York 10461, USA

Giant luminous arcs from gravitational lensing

SIR—The giant luminous arcs observed by Soucail *et al.*¹ and by Linds and Petrosian² have stimulated a number of speculations about their nature. In particular, in a recent News and Views piece, Paczyński proposed that they could be gravitationally lensed images of galaxies³. Although this proposition can be directly tested by measuring the redshifts of the arcs, the measurements are difficult because the arcs are faint.

So I have carried out an immediately available, if less conclusive test, which is to estimate an *a priori* probability for the proposition to be realized. Such an estimate may not be very meaningful for a giant black hole lens, because very little has been established observationally regarding the existence or numbers of such objects. Much more is known about clusters of galaxies, which are an alternative possibility as a suitable lens.

From available cluster statistics, there are about 2,000 possible clusters for such

lensing³ (these are clusters of estimated characteristic central surface mass density above 0.4 g cm^{-2} and situated roughly between redshifts 0.25 and 1).

An arc-like image of a galaxy like that observed can be formed subject to the following two conditions. First, the cluster mass distribution, as projected on the sky, must be close to circular. The degree of the deviation is measured by the reduced shear, γ_c , which is defined in ref. 5. The reduced shear is a combination of the asphericity of the cluster gravitational field and of the additional asymmetry which may possibly be introduced by other massive objects, say other clusters, projected on the sky close to the cluster in question. It also depends on the redshifts of both the lens and the source. The reduced shear determines the angular size of a cone of astroid cross-section within which a point source must reside to produce four or five images. The solid angle inside the astroid is⁵

$$S_c = \frac{3\pi}{2} \gamma_c^2 \theta_c^2,$$

where θ_c is a characteristic deflection angle by the cluster. It can be roughly estimated by replacing the cluster by an isothermal sphere of the same line of sight velocity dispersion, v_l ,

$$\theta_c \approx 4\pi \frac{v_l^2}{c^2} \frac{d_{ls}}{d_{os}}$$

where d_{ls} and d_{os} are the lens-source and observer-source angular diameter distances. The clusters relevant for lensing have unusually high velocity dispersions⁴. For the order of magnitude estimates, $v_l \sim 1,500 \text{ km s}^{-1}$ and $d_{ls}/d_{os} \sim 0.5$, θ_c is $\sim 30''$.

The second condition for producing large arcs is that the lensed galaxy covers a significant portion of the astroid, say a quarter, for the arcs found by Lynds and Petrosian. The sizes of observable parts and the number density of the relevant, high redshift galaxies^{6,7} ($z \sim 0.5-2$ can be used to give a rough estimate of $\sim 2''$ for the angular radii of the observable parts of characteristic (L^*) galaxies and $\sim 100 \text{ arc min}^{-2}$ for their number density).

For the astroid to be of a size comparable to that of a galaxy, the shear parameter must be $\gamma_c \leq 0.03$. The distributions of this shear for clusters are not known, but very rich clusters should be well virialized and extremely asymmetric mass distributions are unlikely to be characteristic. The statistics of multiply ranged quasars⁸ place the characteristic values for the shear to be ~ 0.5 . An estimate on the basis of a uniform distribution between 0 and 1 yields probability $p_\gamma \sim 0.03$ for a shear suitable for producing arcs.

The probability that a galaxy will cover a large part of the astroid is $p_g \sim 0.3$, so that out of the 2,000 clusters about 20 are expected to produce luminous arcs similar

to those found by Lynds and Petrosian.

Such a large number may be inconsistent with the absence of such arcs in cluster samples taken by Gunn and Schneider, as noted by Paczyński. It must, however, be clear that the present number is only an order of magnitude estimate. The main difficulty in achieving a much more precise estimate is that the observational data on mass distributions in clusters of galaxies is not available for a part of clusters in the known complete samples. But this estimate shows that we could have anticipated the discovery of the giant luminous arcs, that Paczyński's proposition is completely realistic, and that it may be possible to find some other arcs around the richest of the high-redshift clusters.

ISRAEL KOVNER

Physics Department,
Weizmann Institute,
Rehovot 76100, Israel

1. Soucail, G., Fort, B., Mellier, Y. & Picat, J.P. *Astron. Astrophys. Lett.* **172**, 414-416 (1987).
2. Lynds, R. & Petrosian, V. *Bull. Am. astr. Soc.* **18**, 1014, (1986).
3. Paczyński, B.P. *Nature* **325**, 572-573 (1987).
4. Kovner, I. *Astrophys. J.* (in the press).
5. Kovner, I. *Astrophys. J.* **316**, 52-69 (1987).
6. Tyson, J.A. *Astron. J.* **92**, 691-699 (1986).
7. Waldrop, M.M. *Science* **234**, 1202 (1986).

Fitness of insecticide resistance

SIR—Clarke and McKenzie¹ and McKenzie *et al.*² have convincingly shown that continued use of diazinon against sheep blowfly, 14 years after resistance to this organophosphate was first detected, has gone on to select modifiers which minimize fluctuating asymmetry and deleterious effects of the resistance allele. These same authors have also shown that blowflies resistant to dieldrin, a cyclodiene, are lacking in modifiers despite high asymmetry and large fitness disadvantages in the absence of dieldrin. But their explanation for the absence of dieldrin-resistance modifiers — that use of dieldrin after resistance evolved was too limited — may be insufficient, for two reasons.

The first is that cyclodiene usage was more extensive than they suggest: γHCH (a cyclodiene-type insecticide known to rapidly select dieldrin-resistant blowflies³) had been widely used against sheep ectoparasites for several years prior to the introduction of dieldrin⁴ and ought to have selected modifiers had they been available. The second reason is that effective modifiers of dieldrin resistance may not exist at all. This opinion stems from the recently discovered biochemical mechanism of dieldrin resistance in cockroaches^{5,6} and work by myself on the fitness and behaviour of dieldrin-resistant mosquitoes. Despite the diversity of test insects used, the conclusions are probably mutually applicable because the dieldrin-resistance mechanism seems to be the same in all species (characteristics held in common include semidominance in

heterozygotes, and parallel cross-resistance spectra and resistance factors)⁷.

The normal toxic action of dieldrin is to bind to receptors on chloride channels of nerves and thereby prevent entry of chloride ions and block transmission of inhibitory impulses⁸. Dieldrin is ineffective against resistant cockroaches because their receptors have become insensitive to this insecticide⁶. The fitness disadvantage of resistance in a cyclodiene-free environment is behavioural: resistant insects are less active and less responsive to stimuli. Evidence for this comes from observations on four species of mosquito from three continents: in every species homozygotes for resistance spent less time searching for hosts and were less responsive to predator movement than heterozygotes and susceptibles (fitness effects were always recessive and never modified by genetic background); in addition female homozygotes were less responsive to oviposition stimuli and male homozygotes had limited mating success.

My conclusion is that dieldrin resistance raises the response threshold of homozygotes to a range of unrelated stimuli; perhaps the mutation to the dieldrin receptor⁶ has increased the permeability of chloride channels, causing hyperinhibition of the nervous system. I suspect that such a far reaching deleterious pleiotropic effect would not easily be neutralized, and that capable modifiers have not yet evolved despite long periods of cyclodiene selection (three of my four species of mosquito were colonized 20 years after cyclodienes were first used).

Insecticide resistance genes give such an enormous advantage in the presence of insecticide that what would otherwise be quite major disadvantages might count for little. While not disputing the importance of time in the selection of modifiers, their existence cannot always be assumed. Before Clarke and McKenzie can pull dieldrin resistance into their argument, they first need to show that modifiers have been selected in areas where cyclodienes are still used and where dieldrin-resistant blowflies are present at high frequency. The Lismore area of New South Wales might be a good place to look⁸.

M.W. ROWLAND

Crop Protection Division,
Rothamsted Experimental Station,
Hertfordshire AL5 2JQ, UK

1. Clarke, G. M. & McKenzie, J. A. *Nature* **325**, 345-346 (1987).
2. McKenzie, J. A., Whitten, M. J. & Adena, M. A. *Heredity* **49**, 1-9 (1982).
3. McKenzie, J. A. & Purvis, A. *Heredity* **53**, 625-634 (1984).
4. McKenzie, J. A. & Whitten, M. J. *Aust. J. biol. Sci.* **37**, 45-52 (1984).
5. Kadous, A. A., Ghiasuddin, S. M., Matsumura, F., Scott, J. G. & Tanaka, K. *Pestic. Biochem. Physiol.* **19**, 157-166 (1983).
6. Tanaka, K., Scott, J. G. & Matsumura, F. *Pestic. Biochem. Physiol.* **22**, 117-127 (1984).
7. Oppenorth, F. J. *Comprehensive Insect Physiology, Biochemistry and Pharmacology*, Vol. 12 (eds Kerkut, G. & Gilbert, L.) 731-773 (Pergamon, Oxford, 1985).
8. McKenzie, J. A. *Aust. J. biol. Sci.* **37**, 367-374 (1984).

The governing influence

David Joravsky

The Communist Party and Soviet Science. By Stephen Fortescue. Macmillan, London/ Johns Hopkins University Press, Baltimore: 1987. Pp.234. £27.50, \$28.50.

SCIENCE is an essentially autonomous enterprise that somehow serves the powers that be, which presents a two-sided puzzle. How is it that scientists, freely investigating whatever provokes their curiosity, are so routinely subservient, so infrequently subversive? One must think back to the age of Darwin and Marx for the last time that claims of scientific knowledge gave serious offence to the dominant ideology, that is, the self-serving beliefs of dominant people. And, on the other side, how is it that the power élite, wilfully pursuing aggrandizement of military force, financial wealth and political domination, is so abundantly served by scientists, with only the occasional murmur against excesses? Voltaire or Diderot, restored to life, would hardly see in this symbiosis the triumph of Enlightenment, considering how determinedly unenlightened the rulers of the twentieth century have been, how much wholesale slaughter their use of science has caused and threatens still, how tame the men of knowledge are, how far from *philosophes* they have declined.

One picks up Stephen Fortescue's book with hope that the two-sided puzzle will be illuminated, if not solved, for the Soviet power élite has been scandalously brutish in its manner of rule, yet it has also been an avid patron of science. Surely this is an extreme test case, which can reveal the relationships that enable men of power and men of knowledge to keep each other's enterprise going. For a while Fortescue seems well launched upon the

political side of the puzzle. He starts out by neatly sketching three favoured "models" of the Soviet system: the party over society as its totalitarian master (old-fashioned version and neo-), the party as central executive in a huge corporate system, and a pluralist model of the party as broker among competing interest groups. His book, he says, will use the party's relations with scientists to see which of those models comes closest to reality. And, true to his word, in the Conclusion he tells what elements of Soviet reality he perceives in each of the three models. But in between the exciting start and the eclectic end the disappointed reader finds little significant analysis of how things actually work between the party and the scientists, how one side manages to rule, the other to generate knowledge, with mutual benefit or at least without much disturbance to either of them.

Of course, there have been some major disturbances in certain fields — genetics, most notably, and all the social sciences — and mutual benefit has not been proportional to the very large investment in scientific education and research. Fortescue repeatedly comments on both problems, but he does not go beyond generalities. He notes, for example, the ups and downs that economics and sociology have endured, but only in a superficial way, without consideration of the issues that have caused political authorities sometimes to approve, sometimes to disapprove, this or that kind of research in

fields that bear directly upon the politicians' business. Science is largely absent from this survey of science and politics, and so too, surprisingly enough, is politics. Organizational detail abounds, with intermittent speculation on its possible effects upon science and politics.

Fortescue does make an extended effort to comprehend "the ideology of science", that is, the mentalities that scientists and party leaders have brought to their interaction, but clarity eludes his analysis. His theoretical concepts are not sufficiently distinct, and his illustrative examples are not sufficiently concrete. More often than not he seems to share the common assumption that ideology is a fantastical opposite of practicality, which is supposed to be self-evident common sense plus the experts' detailed elaboration of self-evident common sense. In this common assumption ideology is characteristic of the Soviet Union's old-fashioned political leaders, while new "technocrats" and scientists make common cause on behalf of practicality. Yet Fortescue also seems intermittently to think of ideology in the anthropological fashion, as the self-serving beliefs that hold a group to some common pattern of behaviour. In that sense, practicality, perhaps science itself, is perceived as a form of ideology, and the analyst of science and politics gets insight into the cooperation and the quarrelling that attend their interaction.

Consider IQ tests, for example. They appealed to Soviet psychologists and to political leaders as a practical way of sorting out schoolchildren and job applicants — until 1936, when the Central Committee suddenly denounced them as an ideological artefact of "bourgeois" science, for they had been mostly denying advancement to workers' and peasants' children. Later on, in less brutal fashion but with analogous reasoning, American school systems cut out mass use of IQ tests, and experts are still arguing about the relative weight of scientific knowledge and political commitment in the process. Concrete instances of that universal sort are conspicuous by their absence in Fortescue's book.

Also absent are instances peculiar to Communist systems. Take land rent, for example, its analysis by economists, its mode of extraction and distribution in the social system. In the course of collectivization Stalin declared that land rent disappeared along with private property in land. But of course different parcels of land still yielded different rates of return, and in 1944 Stalin acknowledged that. With characteristic impudence he blamed Soviet economists for ignoring the obvious, and ordered them to start working out the best way to measure the differences, so that some optimal distribution of land rent could be devised.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

That problem is still far from being solved efficiently, for it involves much more than mathematics. The whole cumbersome system of agricultural pricing and state procurements is involved. If Fortescue had dug into examples of that kind, his fuzzy references to the ups and downs of mathematical economics would have become clear and meaningful, and he would not have had to declare, towards the end of his book: "the intricacies of the politics of these two disciplines [economics and sociology] are, I fear, beyond human understanding" (p.161).

Vnedrenie, the diffusion of improved techniques, is the main problem that Soviet political leaders perceive in science. Fortescue shows that they are inclined to that vulgar notion of science not only by ideological predilection, but also by the obstructions to *vnedrenie* that are built into their economic system. Centralized planning rewards managers for performance of assigned tasks, measured in units of prescribed output, not for taking risks with new techniques. Fortescue argues that party leaders keep trying to get scientists involved in the business of *vnedrenie* as agents of reforming pressure on business managers, who are drawn to conservatism by self-interest. It is a provocative argument, suggesting once again the vicious circularity of so many efforts to reform Soviet behaviour by pressure from above, without reforming the system that shapes the behaviour. Whether the most recent moves for reform are significant departures from that old pattern — some trace it back to Peter the Great or even to Vladimir — Fortescue does not attempt to discover.

The system of personnel selection and promotion that enmeshes Soviet scientists seems designed to maximize intrigue, backbiting and lacklustre hackwork. Fortescue summarizes the picture of that incredible system as it has emerged in angry exposés by exiles and by occasional reports in the Soviet press. He notes that good research is nevertheless encouraged by individual directors of institutes and laboratories who have learned to combine authentic commitment to science with the special administrative skills that the Soviet system requires. But here too his account is vitiated by its generality, its relative lack of vivid examples such as one finds in the writings of Zhores Medvedev. Medvedev's *Soviet Science* remains the best general introduction to the relations between men of power and men of knowledge in the USSR. Fortescue's ruminations on the relative merits of competing models to explain the political system provide a useful supplement. □

David Joravsky is a Professor in the Department of History, Northwestern University, Evanston, Illinois 60201, USA. His book *The Lysenko Affair* was recently re-issued in paperback by University of Chicago Press.

Spinal discord

Keith Stewart Thomson

The Ancestry of the Vertebrates. By R.P.S. Jefferies. *British Museum (Natural History) London/Cambridge University Press, New York: 1986. Pp.376. £50, \$75.*

THE literature is full of ingenious attempts to pin-point an ancestral group for the vertebrates, including that of Willmer who suggested the nemertine worms solely because no one else had. There has, however, been a general consensus that echinoderms are the sister group of chordates and that tunicates are the sister group of amphioxus plus vertebrates. The position of Hemichordata is uncertain. Following Garstang and Berrill, the key stage in the origin of vertebrates has been taken to be the tunicate tadpole larva, in which case direct fossil evidence of vertebrate ancestry is unlikely to be found.

Richard Jefferies has evolved a totally new phylogenetic scheme for the deuterostomes and puts at its centre a controversial argument that members of the paraphyletic fossil group "Calcichordata" (the carpoid echinoderm *Cornuta* and *Mitrata*) can be interpreted as true chordates, variously forming the ancestors of and connecting links among cephalochordates, tunicates and vertebrates. Jefferies bases his argument on cladistic principles and homology of structure. The cladistic argument depends upon the concepts of "stem group" and "crown group". His premise is that, if the echinoderms are the sister group of chordates, then the "fossils traditionally regarded as primitive echinoderms without radial symmetry [i.e. "calcichordates"] should include not only true stem echinoderms, but also stem chordates ..." (p.191). The possibility that they are simply primitive non-radial echinoderms is excluded. This syllogism forces the issue of chordate relationship.

Cornute carpoids have a boot-shaped theca (if echinoderm) or head (if chordate) and a three-part stem, stele or feeding appendage (in various echinoderm interpretations) or tail (chordate). There are openings on one side of the theca/head that Jefferies identifies as left gill openings. The skeleton is calcitic and typically echinodermal. *Ceratocystis* has a hydropore. Mitrates lack gill openings but Jefferies identifies right and left atria (indicating paired internal gills).

Jefferies views the hemichordates as the sister group of all other deuterostomes, but sees in the asymmetry of cornutes a primitive character for his "Dexiothetica" (echinoderms plus chordates) which have "fallen over" to their right side. Therefore the right gills, present in hemichordates, were lost in the common echinoderm-chordate stock and later reinvented (as is

almost literally the case in the ontogeny of amphioxus).

If cornutes and mitrates are chordates, identification of a notochord/dorsal nerve cord/segmental muscle complex is obviously essential. Jefferies's evidence is really an argument, as following (with emphasis added). The tail (if it is a tail) which

seems to be adapted for flexion to left and right... would need muscles and the segmental nature of the skeleton *suggests* that they *might have been* present as muscle blocks... some anti-compressional structure *would be required* in the soft parts of the fore tail and *most probably* it was a notochord in the central axis [p.202].

Then,

evidence from the mitrates... indicates the presence of a dorsal nerve cord and spinal nerve ganglia... which *presumably existed* likewise in cornutes [p.202].

Checking mitrates, however, we find merely that their tail also

...*would need* an anti-compressional device. This was *presumably* the notochord... [p.246].

And the evidence for the dorsal nerve cord and segmental nerves?

A pair of rock-lumps in each segment, corresponding to... paired pits in each ossicle, *Granted* that mitrates are related to cornutes ...[which] *can be seen as chordates*, then it is *reasonable to interpret* these organs by comparison with chordates. *It is then natural* to see the dorsal belt-like organ as the dorsal nerve cord, the roughly cylindrical organ as the notochord, and the paired rock-lumps as represented spinal ganglia... [pp.246-247].

Finally, in cornutes, it is necessary, because there is no other place for it, to place the longitudinal blood vessel of the feeding arm/stem/stele/tail inside the putative notochord. Identification of "trigeminal ganglia" in cornutes, and a "recognisably fish-like brain" in mitrates, follows.

This simply is not enough to make the case. The not uncommon palaeontolog-

New journals review

On 24 September *Nature* will publish the seventh annual review supplement devoted to science journals.

Criteria for inclusion of a journal in the 1987 issue are that:

- (i) the first number appeared, or the journal was retitled, *between June 1985 and May 1986* (the second cut-off date allows at least three issues of a journal to have been published, the minimum number on which a reasonable judgement can be based);
- (ii) it is published at least three times a year;
- (iii) the main language used is English.

Publishers and learned societies are invited to send *four different issues of each suitable periodical, including the first and most recent numbers* (if from outside the United Kingdom, by air mail) to: The Review Editor, *Nature*, 4 Little Essex Street, London WC2R 3LF, England. Subscription details for 1987 (and 1988 if possible) should be included.

ical practice of piling on interpretations, each of which is contingent on some previous part of the argument and fails therefore to test the original premise, is particularly weak when the group in question has no living representative. Here, no indisputably chordate character is available to serve as a foundation for argument, while critics such as Philip and Ubaghs readily interpret the same structures as echinodermal. Perhaps in the end, to quote again:

A reader can only despair.... When equally eminent workers, starting from the same data, reach mutually contradictory conclusions, it might seem that all phylogenetic reconstruction is vain [p.350].

But all is far from vain, especially if the "calcichordate" argument is isolated from Jefferies's other contributions — his new

phylogenetic hypothesis for the living deuterostomes and his challenge to the story-telling that is inherent in the various schemes involving the tunicate tadpole. Jefferies has a masterly command of the anatomical literature, especially German and Russian, and the half of the book that reviews and analyses chordate structure and relationships contains many useful new insights. The general hypothesis can be tested against others based on biochemistry or development, but the "calcichordate" hypothesis cannot. Failing that, Cornuta and Mitrata seem to remain echinoderms and, like Wordsworth's "primrose by a river's brim", nothing more. □

Keith Stewart Thomson is a Professor in the Department of Biology, Peabody Museum of Natural History, Yale University, New Haven, Connecticut 06511, USA.

Outside opinion

Andrew Hill

The Red Ape: Orang-utans & Human Origins. By Jeffrey H. Schwartz. Houghton Mifflin, New York/Elm Tree, London: 1987. Pp. 266. \$18.95, £12.95.

OUR closest relatives among living animals are generally taken to be the African apes, that is, the gorilla and the chimpanzee. Charles Darwin held this view, though not so firmly as is popularly believed, and since then most people have concurred with him. The German zoologist Ernst Haeckel, Darwin's follower in both time and spirit, believed otherwise and suggested that the Asian orang — the red ape of this book's title — is the closest to us. Jeffrey Schwartz has resurrected this viewpoint and during the past few years has been heckling the palaeo-anthropological establishment from the sidelines.

Schwartz's initial forays were published in professional journals, and as well as meeting with surprise, not to say astonishment, also caused the realization that perhaps some of the assumptions on which received notions were based might benefit from closer scrutiny. In this sense his position has proved valuable. If anything, however, re-examination has only served to confirm the orthodoxy in the minds of those concerned.

The problem for Schwartz is that there are so many independent lines of evidence, among them palaeoanthropology, comparative anatomy, molecular systematics and cytogenetics, which support the conventional opinion. Work in each of these fields and others all points to the same conclusion — a strong phylogenetic proximity between African apes and humans. A few excuses for thinking other-

wise can be found in each area — and Schwartz finds them — but to invoke contrary circumstances to explain them all is a little too much.

So, if the evidence really does give the same indications, where is the scope for argument? Well, animals can be similar for different reasons. It has long been known that they can be similar because they are parallel or convergent in evolution, relatively unrelated species showing analogous adaptations to similar circumstances and needs. One of the main contributions of cladistic analysis has been to

derived features? There are methods for assessing the polarity of characters, for allocating similarity to primitive plesiomorphies or to derived apomorphies, but despite the assertions of some committed cladists it often remains difficult to be certain. This disturbing fact is employed centrally in *The Red Ape*.

Schwartz implies that the current view has developed from a misinterpretation of the similarity data, a consensus accepted uncritically by most palaeoanthropologists mainly because of their unquestioning assimilation of tradition. Suggestions of such cultural relativism are popular at the moment, the awareness of it being itself one of the culturally relativistic factors we have to contend with in evaluating ideas. But other scientists may believe this stuff not because they are victims of the myopia of history and their social circumstance, but simply because on more objective grounds they think it to be correct. Whether or not Schwartz turns out to be right in the long run, I still think the present evidence contradicts his view.

The Red Ape is entertaining and plausibly argued. Certainly, there's little wrong in holding or promulgating eccentric ideas — most good ideas were eccentric once — and whether or not you believe Schwartz's basic premise his book contains much of interest. There is a great deal of historical information about orangs and about ideas of phylogeny that will be of value to professionals and will fascinate the general reader. The only danger of the book is that

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Family fracas — the orang-utan, Jeffrey Schwartz's contender for our closest living relative.

make explicit the point that even homologous similarity (similarity due to relationship) can be of two kinds. One is useful in reconstructing phylogeny the other is not. If animals are similar only because they share the primitive character state of some remote past ancestor, then this provides little evidence about their present relationship. On the other hand if they are similar because they share a more recently derived state, then that is good information about relatedness. This is what the argument is about. Which, of the various similarities among the apes and humans are primitive retentions, and which are

people who have the inclination or time to read only a single account of human evolution will choose this one, and will take its contents to be representative of a large body of professional thought — which they are not.

In Huxley, Darwin had his bulldog to promote and support his theory. As he himself more or less admits in the book, Schwartz has yet to find as much as a poodle. In the dogged pursuit of his idea, he is barking up the wrong family tree. □

Andrew Hill is in the Department of Anthropology, Yale University, PO Box 2114 Yale Station, Connecticut 06520, USA.

Deep into the past

Robert Temple

Medicine in China. By Paul U. Unschuld with translations and annotations. *University of California Press: 1985–1986. Vol. 1 A History of Ideas pp. 423, \$45, £35.95. Vol. 2 A History of Pharmaceutics pp. 367, \$65, £51.95. Vol. 3 Nan-Ching: The Classic of Difficult Issues pp. 760, \$69.95, £55.95.*

In this series of books, Paul Unschuld has accomplished monumental labours of translation and annotation. He has a fine historical sense — as for instance when he shows how medicine and politics were inextricably intertwined in second century China during the Yellow Turbans Rebellion — and a colourful and easy style which makes these somewhat esoteric subjects more accessible. But his scholarship, though it can be deep, is not wide. He speaks of his own work as “a completely new approach to Chinese history”, which it probably is. It was all the more incumbent on him, therefore, to take more care. In attempting to survey the entire history of Chinese medicine, he has overlooked works on the subject by Fan Hsing-Chun, Sung Ta-Jen, and Wang Chi-Min and Wu-Lien Te, as well as important articles by Ch'en Chih, Hu Hou-Hsüan, Yü Yün-Hsiu and R.F. Bridgman. Worse, he cites no work of Joseph Needham more recent than 1980, ignoring several important articles as well as the crucial volume of *Science and Civilisation in China* on physiological alchemy published in 1983.

On the few occasions when Unschuld mentions Needham, he takes issue with him. For instance, Unschuld insists that acupuncture was not practised in China before the first century BC, whereas Needham gives evidence indicating a date at least four or five centuries earlier. Actually, the hundreds of “hairpins” found in Shang Dynasty tombs may be acupuncture needles, dating the practice to the fourteenth century BC or earlier. But Unschuld dismisses such possibilities.

The finest of these three volumes is the second, *A History of Pharmaceutics*, which gives a satisfying survey of Chinese pharmaceutical and herbal writings from the earliest to modern times. Here Unschuld achieves great success, made all the more attractive by the large number of early woodcuts of plants, animals and minerals, reproduced for the first time in the West. Many important texts are translated in this volume, and they are linked together by a coherent and readable commentary. Unschuld is at his best here, with only minor annoyances such as the European translations of Li Shih-chen, China's greatest pharmaceutical writer, being omitted from the bibliography.

Volume 3 is a translation of one of the

principal Chinese medical classics, the *Nan-Ching*, dating from the first century AD. This can only be described as a labour of love, because the work is rather boring to anyone not interested in Chinese medical history. But its availability in a sound philological translation is a great help to scholars, and Unschuld has done a great service which will not go unappreciated. There is no index to the translated text itself, which may be a hindrance for some scholars, especially as the *Nan-Ching* is enormously long. Unschuld has thoughtfully also provided translation of commentaries on this work by 20 Chinese and Japanese authors over the past 17 centuries. Indeed, his familiarity with Japanese material is a particularly strong

point, in view of the fact that so much to do with the history of Chinese medicine was lost in China itself and preserved only in Japan.

Future volumes in the series are said to be in preparation, though their specific subjects are not identified. If Unschuld could broaden his range of scholarship and be more careful with citations, his continuing contributions in this field would take on a value proportional to his enormous labours. □

Robert Temple, Hill House, Sutton Mallet, Bridgwater, Somerset TA7 9AW, UK, is a historian and science writer. He is author of *China: Land of Discovery and Invention (in the United States entitled The Genius of China)* recently published by Patrick Stephens/Simon & Schuster.

Analytical answers

Nick Walsh

A Handbook of Silicate Rock Analysis. By P. J. Potts. *Blackie, Glasgow/Chapman & Hall, New York: 1987. Pp. 622. £128, \$175.*

MANY books on silicate analysis have appeared in recent years, but this monumental contribution is different to anything produced so far. It is an attempt to provide a comprehensive account of all the main analytical techniques now in common use for silicate rocks. To undertake this in a multi-author work would be ambitious, but for one author to write such a book is indeed a remarkable achievement. The resulting volume is clear, concise and authoritative, and the discussions of the techniques covered benefit from the uniformity of style. The scope is such that the book should be useful not only to those directly involved in the day-to-day analysis of geological samples but to many other workers as well.

Dr Potts's aim has been to evaluate all of the techniques now used for silicate analysis. Thus the classical (gravimetric) methods are not overlooked, and the “rapid” methods of analysis, developed some 20 years ago (using largely colorimetric/photometric techniques), are presented in adequate detail. Description of modern instrumental methods of elemental analysis takes up most of the book, and there are excellent chapters on atomic absorption, inductively coupled plasma spectrometry and X-ray fluorescence spectrometry (wavelength and energy dispersive). Each of these chapters gives full details of both principles and practice, and with their references they will come to be seen as definitive. The coverage can hardly be faulted, although detailed analytical recipes are not usually given.

A difficulty for any author of a book of this nature is deciding what to include and

what to exclude — it is impossible to accommodate all prejudices. Thus purists may quibble with the appearance of such a detailed treatment of electron probe analysis in a book dealing with rock analysis. This technique is primarily used in the analysis of minerals and its application to rocks is, in the author's word “interesting”. Most readers, however, will surely welcome its inclusion, and also that of the short but helpful accounts of “Other Microbeam and Surface Analysis Techniques” (ion probe, laser microprobe, particle-induced X-ray emission, electron spectroscopy for chemical analysis and transmission electron microscopy) that follow. Neutron activation analysis and mass spectrometry (thermal ionization and gas source) are also well represented. Some researchers would probably have preferred to have an even more detailed discussion of isotope dilution methods, in view of their high precision and accuracy. Others may regard the treatment as more than adequate, given the time and cost involved in such analysis. Overall, the balance between description of the techniques themselves, and evaluation of their potential advantages and disadvantages, is the best that could have been attained.

Many other useful topics are included. There are, for example, chapters on ion-selective electrodes, ion-exchange separation techniques, gold and platinum group analysis, spark source and inductively coupled plasma mass spectrometry. Of real use to practising analysts are the chapters “Concepts in Analytical Chemistry” and “Optical Spectrometry: Principles and Instrumentation”, which provide answers to many fundamental questions.

This is a fine volume, well written and beautifully presented. It is by far the most important book on silicate analysis published in recent years, and given the amount of information included well justifies its apparently hefty price. □

Nick Walsh is a Lecturer in the Department of Geology, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, UK.

Engineering a molecular nightmare

Erwin Chargaff

The application of new reproductive technologies carries all manner of threats to the dignity of human life. Should not scientists show some restraint?

It is often said that progress — social, economic, scientific and even cultural — is driven by technology. An alternative view can also be heard: that requirements, demands, urges, arising in an unexplained manner, generate the techniques suitable for satisfying them. In other words: either "Where there is a way there is a will" or "Where there is a will there is a way".

The semi-industrial production of babies seems to belong to the first category: the demand was less overwhelming than the desire on the part of the scientists to test their newly developed techniques. The experimental babies produced were more of a by-product, helping to finance, through their "parents" and also through public funds, the progress in fetal research. Such a term as "reproductive technology" is, actually, a portmanteau designation, covering many activities, some of them perhaps not very pretty. It comprises all sorts of fertility research, fetal pathology, embryo transplantation and so on. To me it would seem to cover even such heart-rending cases as that reported in *Nature* of 5 June 1986 (p.553) in which a pregnant schizophrenic woman was persuaded or forced to consent to an abortion so that the physicians could examine the fetus in great detail. This is, however, not the direction that the present article will take.

Objections

What I want to do here is to formulate my personal objections to several types of research now being performed in the field of human reproduction. I do this as an old biochemist whose previous areas of interest, mainly in the chemistry of the cell, are tangential to the subject of my present remarks. But, even more, I do it as a man wishing to protest against the continuous decay of a society that is being misled by roseate prospects, in order to screen undesirable deeds. I believe it is time to contrast the maxim, generally accepted by our scientific civilization, "It is the end that sanctifies the means" with another one: "It is the means that diabolize the end".

It ought to be emphasized that mine is not the voice of the expert. That is, perhaps, not as much of a stricture as it would seem. We have been overdoing our admiration of expertise. Judges, after all, are not usually specialists in breaking and entering. Expertise often paralyzes com-

mon sense. What it is good for is to know how to get certain things done, but it cannot decide whether these are morally desirable or objectionable. It is, in fact, blind to such discriminations.

When, as a consequence of one or the other of the techniques of reproductive technology now being elaborated, a baby is born we shall say that the intervention has been successful. Is that not too hasty a conclusion? The fate of a human being begins with conception, but it does not end with birth. How many years will have to elapse — ten, fifteen, twenty? — before we may claim success? Is it completely out of the question that during normal conception factors play a role that are not duplicated in the synthetic milieu in which *in vitro* fertilization takes place? (I am not thinking of mystical influences but of good solid chemistry.) In manipulating processes worked out by nature in the wisdom of millions of years one must be aware of the danger that our shortcuts may carry a bleeding edge.

In normal conception the female egg is confronted with a very large number of viable spermatozoa. Fertilization could then appear as a purely random event, or, alternatively, it could be looked upon as a process in which the egg selects the sperm cell with which to fuse. This is not a metaphysical quibble, though it may be an as yet unanswerable question: are we sure that the nucleotide sequence of the DNA contained in two human spermatozoa is completely identical, although both cells carry all the genes required for the definition of the species? A large part of human DNA is unassigned genetically. The biologists speak of "the human genome", they speak of "nonsense DNA", but maybe it is they who talk nonsense. It is not at all unlikely that those uncommitted portions of DNA carry functions that cannot be, or have not yet been, foreseen.

We do not know what life is, and yet we manipulate it as if it were an inorganic salt solution. Adding sodium chloride to an excess of a silver nitrate solution, we assume correctly that there is no difference between the silver atoms precipitated as chloride and those remaining in solution. But in effecting fertilization, do we not stick our clumsy fingers into the incredibly fine web of human fate? We select for a child that might otherwise have never been born.

The "pre-embryo" is a designation that

appears to me entirely unjustified. I fear that it has merely an alibi function. The life of the embryo begins with the fertilized egg in which all the potentialities of the organism reside. The attempt to determine, by scientific means, the stage at which what for times immemorial had been called the human soul makes its appearance, is ridiculous. The setting of a calendar date serves only as a permit for the performance of experiments that normal reverence before human life would have outlawed; experiments that until a few years ago would, in fact, have been unthinkable.

Frozen embryos

There now exist techniques for the storage of viable embryos in the frozen state. Such embryos, it has been shown (though I do not know in what percentage of cases), are capable of being transferred, after thawing, to the uterus of the "mother", that is the donor of the egg, or to a hired uterus, the "surrogate mother". As babies have been delivered alive in such a fashion, success is claimed. But rapid freezing of tissues to a very low temperature is not always an innocuous process, so far as macromolecules, for instance nucleic acids, are concerned. The continuously changing electrolyte concentration, as freezing proceeds, and the removal of structurally bound water in the form of ice are not always harmless.

Although the proof of a baby — unlike that of a pudding — does not rest on its consumability, and in the case of man not even on its being born, efforts towards proving that what is good for vegetables is also good for human beings are being continued in many places. The life of man is, however, an unrepeatable experiment: no controls, no placebos. Although the expectations facing, and also checking, a breeder of men must be much higher than those restraining the animal breeder, there will be no one around to do the reckoning and to settle the accounts when the time has come. Human husbandry is so new a profession that society has not yet learned how to protect itself.

Ethics is a branch of philosophy and should be left there. Until recently one would not have expected it to be applicable to pure scientific research. The ethics of chemistry or of geology would seem to make no more sense than the ethics of chess or of violin playing. Nor is "unethi-

cal" a synonym of "criminal". When an industry contaminates a river with cyanides, it does not act unethically, it acts criminally. Only in the recent past, concurrently with the emergence of the new profession of ethicist, have certain direct applications of scientific research begun to be classified as ethical or unethical — qualifications indicating inadequacies of the criminal code.

Thus I find in a recent issue of *Science* (19 September 1986, p.1,255) an article entitled "Ethical Guidelines Proposed for Reproductive Technology". This article, written by Gina Kolata, summarizes a set of recommendations adopted by the American Fertility Society. A gardening manual written by the goatist of goats could not be more permissive. Four categories are distinguished: (1) Ethically acceptable; (2) suitable for clinical trials; (3) suitable for clinical investigations; (4) ethically unacceptable. The first three comprise practically everything that is being done at present. Ethically unacceptable are, however, the patenting of medical procedures and the use of surrogate motherhood for non-medical reasons.

How unselfish scientists are, with their eyes elevated to the empyrean of knowledge and truth, may be gathered from the fact that nowhere — at least in the summary available to me — does it seem to be considered that during all those operations much money must be changing hands, a circumstance that often changes the ethical aroma.

The confusion besetting society, when dealing with some of the fruits of present-day scientific research, is not surprising. It was the wont to rank science, until far into our century, among the highest and purest pursuits of mankind. Science was the never-ending search for truth about nature, a quest that would help us understand the workings of our world. That era has ended, I believe, with the splitting of the atomic nucleus, with the manipulation of the cellular nucleus, with the ability to modify the hereditary apparatus. A new era has begun: science is now the craft of the manipulation, modification, substitution and deflection of the forces of nature. This does not, of course, apply to all scientific disciplines, but very much so to the topic of my present discussion, a form of applied biology that is in great danger of leading to various brutalities which society, once awakened, may not wish to tolerate.

At this point a caveat must be inserted: my words should not be taken to mean that I am erecting a dichotomy of a guileless society facing an evil science. Science grows, of course, out of society and could not exist without its support. It is, in fact, through this support that society, in the form of its representatives, can exercise control. But controls usually come too late, for the people, like the proverbial

generals, always conduct the preceding, not the present war. To all of us, grown up in an atmosphere of freedom of thought and expression, it seems unthinkable that there can be excesses of scientific research. Reproductive technology, in the form in which it is described to me, appears to represent one such example.

I am not equipped to conduct a philosophical discussion of what is meant by human destiny or fate; a discussion that would also deal with the question to what extent the "healing arts" of the past also constituted a struggle against fate. I should think that the very inefficacy of those arts obviated that question. The boundaries of what was healable were pushed back slowly and gradually. Public

"If medicine persists in the path it seems to have chosen at present, we shall soon hear that its great mission actually is the abolition of death . . ."

health precautions and antibiotics dealt effectively with pathogenic invaders of the body, but the body itself remained the vehicle of its destiny. That mysterious designation included all sorts of afflictions, such as a failing heart or the childlessness of a couple. Fate had to be borne, just as did, and still does, the advent of that great drawer of sums, death. If medicine persists in the path it seems to have chosen at present, we shall soon hear that its great mission actually is the abolition of death, a prospect that even the most sanguine of optimists would hesitate to extend to more than himself.

If one of the academies of science, which in the past often announced prize competitions for an essay answering a given question, had been bizarre enough to propose as the topic "How to Render Sexual Intercourse Unnecessary Without Interfering with the Propagation of Mankind?", the present innovative techniques of reproduction could well have been worthy of the medal. That such a question could not have been proposed is not only due to the backwardness of the old sciences, but also to the fact that they were much less meddlesome than they are now, preferring to avoid the inevitable. Whether their ethical sense was more developed than it is now I do not know; but I should say that their aesthetic feelings were: they shied away from the unseemly, the ugly. The manufacture of human embryos with the only purpose of crushing and extracting them is something that I cannot imagine in the hands of a von Baer or a Virchow.

Had I anything to say I would certainly proscribe the production of human embryos for experimental purposes. Scientific curiosity is not an unbounded

good, although I will not deny that, for instance the problem of cell differentiation during embryonic growth is of great interest. Restraint in asking necessary questions is one of the sacrifices that even the scientist ought to be willing to make to human dignity. There is so much we do not yet know; let us start with something else.

I should similarly frown upon the use of surrogate mothers, especially for a fee. It is true, Chapter 30 of *Genesis* gives a colourful description of a very confusing situation, but we shall have to leave it to that ancient desert people to come to terms with their conscience, although in that celebrated account one is dealing really with a form of adoption in a primitive polygamous community.

The giant strides that science — especially in its applications — has made in our time are composed of an infinity of tiny steps, each one appearing to the public as harmless or even beneficial. Helping a few couples condemned to childlessness towards getting a child may strike the obstetrical cytologist as such a laudable step. But we can already see the beginning of human husbandry, of industrial breeding factories. And who can hinder the mass production and industrial exploitation of human embryos, the emergence of a new branch of biotechnology? Who can deny the scientific interest attaching to the production of chimaeras, to the study of human embryonic growth in an animal uterus? Well, I think society could prevent all that and more, but I fear it will not. What I see coming is a gigantic slaughterhouse, a molecular Auschwitz, in which valuable enzymes, hormones and so on will be extracted instead of gold teeth.

"Nothing will remain unavenged" we are being assured by a celebrated line of the *Dies Irae*. That time not yet having come, however, and people being what they are, I see only two ways of preventing even worse from happening, and that is to moderate two human traits, cupidity and ambition. If the money nexus were cut so that all that now makes up reproductive technology would have to be performed without compensation, as truly samaritan deeds, if anonymity were enforced, so that all the remarkable achievements of that new applied science would have to be published without authors' names, almost all regrettable excesses would be avoided. The ridiculously Utopian character of these recommendations should be taken as a measure of my hopelessness. □

Erwin Chargaff, 350 Central Park West, New York, New York 10025, USA, is Emeritus Professor of Biochemistry of the College of Physicians and Surgeons, Columbia University. An Austrian by birth, he was forced to leave Europe by the rise of the Nazis. In work at Columbia in the late 1940s he demonstrated that the ratios of the bases in DNA are constant, a result that was crucial to Watson and Crick in formulating their model of DNA.

Geological terrains and crater frequencies on Ariel

J. B. Plescia

Branch of Astrology, US Geological Survey, 2255 North Gemini Drive, Flagstaff, Arizona 86001, USA

The southern hemisphere of Ariel, a satellite of Uranus, can be divided into several terrain types. Data on the size-frequency distribution of craters for those different terrain types indicate that these terrains formed over a relatively short period of time. Much information on Ariel's geological history can be gained from these data.

ARIEL was imaged by Voyager 2 during its encounter with the uranian system in January 1986¹. Because of the obliquity of the uranian system (about 98°) and the current orientation of the rotational axis with respect to the Sun, Voyager 2 viewed only the southern hemisphere of Ariel (Fig. 1). Only about 35% of the total surface area of the satellite (65% of the southern hemisphere) was observed at a resolution and under viewing conditions appropriate for geological study.

Before the Voyager encounter very little was known about the uranian satellites, even their size and mass were subject to considerable uncertainty^{2,3}. Though limited, the Voyager data have allowed a general characterization of the geology of the uranian satellites. Here we briefly describe the geological terrains recognized on Ariel, present data on the size-frequency distribution of impact craters for those terrains and discuss the implications of these data on the satellite's geological history. The nomenclature used to describe the geographical features is preliminary (H. Masursky, personal communication).

Geology

The surface of Ariel (Fig. 1) can be divided into three terrain types: cratered terrain, ridges terrain and plains (Fig. 2). Differentiation of these units is based on surface morphology and apparent crater frequency. Areas mapped here as cratered and ridged terrains are divisions or subdivisions of the "cratered terrain" of Smith *et al.*¹, whereas plains correspond directly to their "smooth terrain". The full extent of each of the terrains is unknown as each extends into the unimaged northern hemisphere or into areas where little surface detail is discernible.

Several types of tectonic features are observed on Ariel, the most obvious of which are the grabens (Fig. 1). Grabens are 15 to 50 km wide and have varied morphology. Some are degraded, such as Kewpie, Brownie and Pixie Chasmata, whereas others, such as Kachina Chasmata, appear more pristine. Isolated scarps and lineaments also occur, principally within the cratered terrain. Most tectonic features have a north-east trend; only a few have north-west trends.

Cratered Terrain. Cratered terrain (Fig. 2) is the most areally extensive unit. It is characterized by impact craters superposed on a rolling surface which locally exhibits narrow (<3 km wide) sinuous ridges. Grabens and minor scarps cut the surface and albedo variations give the surface a blotchy appearance. Some parts of the terrain appear to have muted topography and the impact craters in these areas appear to be degraded. Such muting does not appear to be the result of high illumination angles which can decrease the amount of surface detail. Much of the muted topography is observed under lighting conditions (35–90° incidence angle) similar to that where crisp topography is observed.

Ridged Terrain. Separating the cratered terrain into polygons and locally bounding the cratered terrain is a unit referred to as ridged terrain (Fig. 2). Ridged terrain is characterized by east- or north-east-trending bands, 25 to 70 km wide, in which are parallel ridges and troughs. Spacing between ridges is typically 10 to 35 km. Individual ridges extend for distances of 100 to

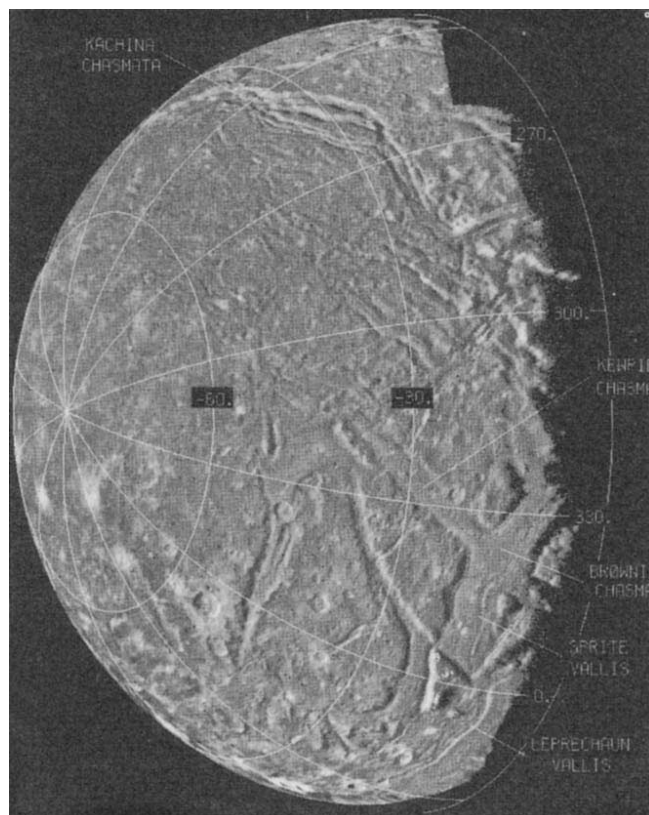


Fig. 1 Composite mosaic of Ariel, centred on lat. -30°, long. 315°. Resolution is 1.2 km per pixel.

200 km, although the bands extend for many hundreds of kilometres. Superficially, this terrain resembles the "ridged plains" of Enceladus⁴. Although all of the terrains exhibit faults, the ridged terrain is mapped separately because of the number and consistent patterns of the faults.

Bands of ridged terrain are, in part, continuous with the grabens that cut the cratered terrain (Figs 1 and 2); this continuity suggests that the ridged terrain may have formed under a tensional stress. But whether the ridged terrain represents a different type of crustal response to the tensional stresses than that which resulted in the grabens of the cratered terrain, a later modification of the grabens or some other tectonic process is uncertain.

Plains. Smooth areas having relatively few craters are termed plains (Fig. 2). Plains occur in topographically low areas: graben floors and other irregular depressions (such as at latitude -45°, longitude 325°). Different areas of plains may have been formed by different processes or be composed of different material. In a few locations the edges of the plains are marked by a lobate pattern. The nature of some of these margins is equivocal, but

many appear similar to the margins of lava flows, which suggests that the plains material was emplaced as flows.

A common feature of the plains observed within the grabens is a linear to sinuous medial trough. Material adjacent to the trough (Sprite Valles) in Brownie Chasma (Fig. 1) is slightly brighter and smoother than material farther from the trough. Similar relations are observed along the branching troughs that occur on the expanse of plains in the graben in which Leprechaun Vallis occurs (Fig. 1). This brighter, smoother material has flowed laterally from the trough, covering older, darker plains, again suggestive that the plains were emplaced as flows. Troughs are also observed along the edge of the plains at the base of the graben walls, but whether these troughs acted as vents is unclear.

Crater morphology

Recognizable craters on Ariel range in diameter from ~85 km to the limit of resolution (~2 km per pixel), although those having diameters <15 km were poorly resolved. Craters typically exhibit flat floors and a central peak; a few of the central peaks are elongated in a north-west direction. Many of Ariel's craters are polygonal in outline; a few are particularly so, with the linear parts of the crater rims in a direction north-east or north-west parallel to the major structural trends. Polygonal craters are observed on many different bodies and have been interpreted to reflect the influence of pre-existing crustal structure on the crater formation process.

Several features are observed on Ariel that have circular outlines ~100 km in diameter. These features are subtle and difficult to recognize as they occur at high latitudes, in the areas of high illumination angle where surface detail is poorly resolved. Their morphology is suggestive of that of the highly modified impact structures observed on Ganymede⁵, the palimpsests, whose topography has relaxed due to the viscous creep of the ice crust.

Ejecta deposits are associated with only a few craters; these deposits have albedos either similar to or higher than that of the terrain on which they occur. Low-albedo deposits (for example, dark-ray craters) have not been observed. Craters having bright ejecta deposits appear to be more common at high southerly latitudes, the region of near-vertical illumination. High illumination enhances the visibility of bright ejecta, hence their increased frequency may be an artefact of the lighting rather than a real change in frequency.

Crater frequencies

Impact crater size-frequency distributions were compiled for the three different terrains to determine their relative ages; the data are listed in Table 1. Initially, each terrain was subdivided into several counting areas to determine the degree, if any, of internal age variation. However, due to the small size of these subdivided counting areas ($5\text{--}55 \times 10^3 \text{ km}^2$) the statistical uncer-

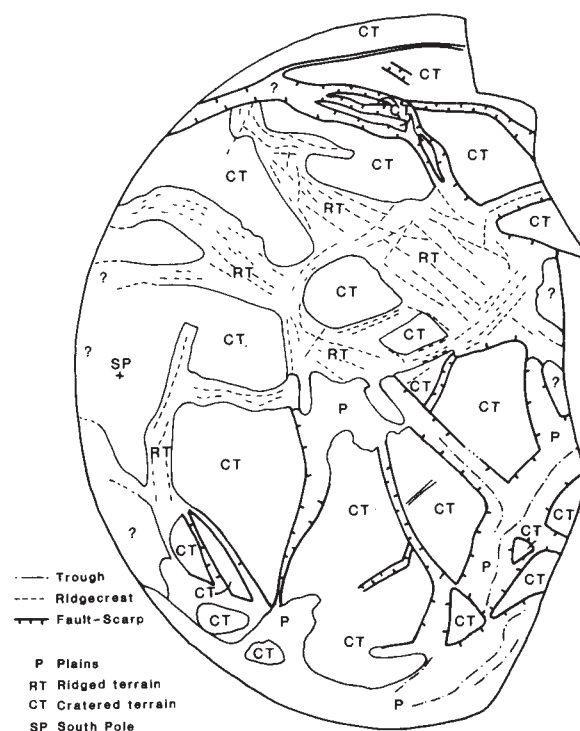


Fig. 2 Sketch map of Ariel showing the distribution of different terrain types. Queries denoted uncertainty as to terrain type.

tainties were large and only substantial age variations would have been recognizable. As no such variations were recognized, the data were combined into a few representative distributions.

Despite the varied morphology of the surface of Ariel, large differences in crater frequency are not observed. The range of crater frequencies between the most heavily cratered and the least cratered regions is less than a factor of 4 (Table 1). This is considerably less than the order of magnitude variations in crater frequency observed on Miranda^{1,6}.

Figure 3 illustrates the cumulative size-frequency distributions for major expanses of the cratered terrain. The similarity of the three distributions indicates that there are no major differences in the age of the terrain. Differences in crater frequency at small (5 to 10 km) diameters (Table 1) may reflect the effects of local resurfacing processes. The distribution for the ridged terrain (not shown) is identical to that for the cratered terrain; this indicates that the ridged terrain is of similar age to the cratered terrain.

Plains units display variable crater frequencies (Fig. 4). The frequencies are similar at diameters >20 km, but the statistical

Table 1 Crater frequencies for Ariel terrains

Region	Frequency of craters $\geq D/(10^6 \text{ km}^2)$					D_{max}^*	Area [†]
	2	5	10	20	30		
Cratered terrain (lat. -15° , long. 0°)	(22,967 $\pm$ 586) [‡]	2,285 $\pm$ 185	358 $\pm$ 73	75 $\pm$ 33	(19 $\pm$ 17)	29	66,962
Cratered terrain (lat. -45° , long. 270°)	(29,663 $\pm$ 582)	2,497 $\pm$ 169	376 $\pm$ 66	62 $\pm$ 27	33 $\pm$ 19	69	87,706
Cratered terrain (lat. -15° , long. 255°)	(63,908 $\pm$ 1302)	4,007 $\pm$ 326	849 $\pm$ 150	69 $\pm$ 43	(34 $\pm$ 30)	29	37,684
Ridged terrain	(37,106 $\pm$ 889)	2,641 $\pm$ 237	547 $\pm$ 108	23 $\pm$ 22	(18 $\pm$ 20)	20	46,994
Plains (lat. -45° , long. 320°)	(24,484 $\pm$ 1510)	1,955 $\pm$ 427	162 $\pm$ 123	(46 $\pm$ 65)	(15 $\pm$ 37)	13	10,740
Plains (lat. -15° , long. 30°)	(11,365 $\pm$ 548)	1,311 $\pm$ 186	198 $\pm$ 72	87 $\pm$ 48	34 $\pm$ 30	33	37,893
Plains (lat. -15° , long. 345°)	(13,769 $\pm$ 905)	1,873 $\pm$ 33	725 $\pm$ 208	106 $\pm$ 79	(68 $\pm$ 64)	24	16,819

* D_{max} denotes the diameter of the largest crater in the sampling area.

[†] Area denotes the area of the sampling area in km^2 .

[‡] Numbers in parentheses indicate extrapolations of the data. This is necessary at small diameters due to resolution limits and at large diameters due to a lack of craters.

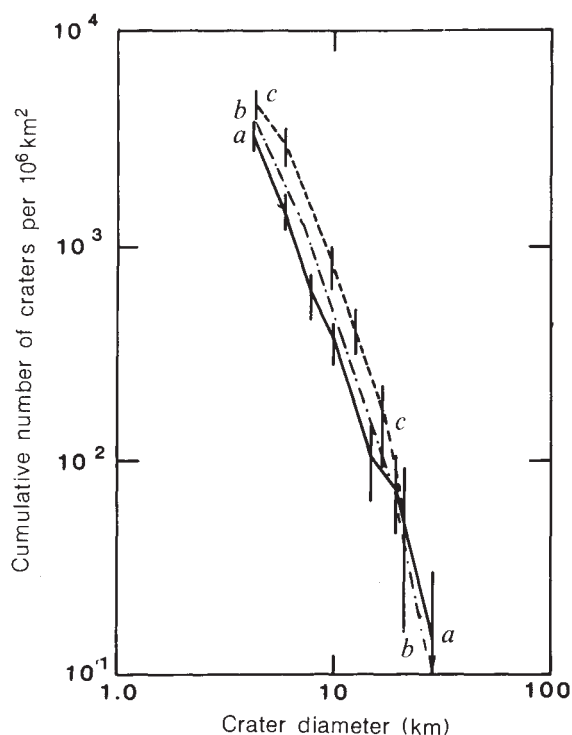


Fig. 3 Cumulative size-frequency distribution for three areas of cratered terrain. Curve *a*, terrain at lat. -35° , long 0° ; curve *b*, terrain at lat. -45° , long 270° ; curve *c*, terrain at lat. -15° , long 255° .

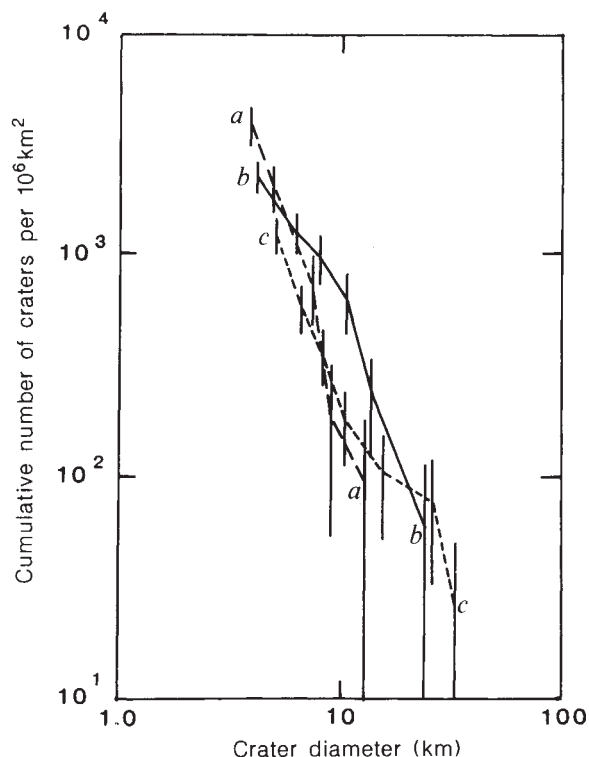


Fig. 4 Cumulative size-frequency distributions for three areas of plains. Curve *a*, region near lat. -45° , long. 320° ; curve *b*, graben fill along lat. -20° , between long. 330° and 0° ; curve *c*, the area at lat. -15° , long. 30° .

uncertainties are large at these diameters because of the low number of craters. However, the statistically significant differences at diameters < 20 km indicate variable ages. The distributions for the plains are structured rather than linear distributions suggesting that the plains-forming events were not simply the complete burial of the pre-existing surface during a geologically short interval of time. Material was probably deposited in layers of variable thickness over an extended period of time.

Discussion

The relative paucity and degraded morphology of large-diameter impact craters and the overall low crater frequencies (the largest impact crater observed is 85 km in diameter), relative to Oberon and Umbriel⁸, suggest that Ariel has been completely resurfaced since it accreted. The heavily cratered surface that presumably formed during the accretion of Ariel, that is, one that resembled that of Umbriel or Oberon, has been replaced. The destruction of the accretionary surface could have resulted from some type of wholesale crustal overturning, perhaps related to interior convection, or possibly from the complete burial of the original surface by outpouring of material from the interior. The exact mechanism by which such replacements occur on any satellites is not well understood.

The differences in crater frequencies at small diameters for the cratered terrain may be indicative of local resurfacing. Material that partly filled the grabens may also have been deposited on the cratered terrain, burying craters and thereby reducing the observed crater frequency. Unfortunately, the resolution of the images is insufficient to reveal the details of any resurfacing processes that may have occurred in the cratered terrain.

The ridged terrain has crater frequencies indistinguishable at all diameters from the cratered terrain, hence its stratigraphic position is ambiguous on the basis of crater frequency. However, the geometrical relation between the ridged and cratered terrains

can be used to determine their relative ages. The ridged terrain separates parts of the cratered terrain into polygons in manner similar to the way the grooved terrain on Ganymede breaks up the dark cratered terrain. This relation indicates that the ridged terrain is younger than the cratered terrain.

The extensive network of grabens indicates that Ariel has experienced global tensional stresses. The timing of graben formation is, in part, constrained by the observed stratigraphy. Grabens cut the cratered terrain; therefore, graben development postdates the formation of those surfaces. At least some grabens (the floor of Kachina Chasmata was not observed) are partly filled by plains material, which indicates that some of the grabens developed before or during the period of plains formation. Later graben development, after plains formation had ended, is also possible.

Ariel's geological history can be summarized in the following sequence. After accretion, a period of global, endogenically driven resurfacing occurred that destroyed the original accretionary crust. During that time, global development of the cratered terrain probably occurred. Subsequently, the cratered terrain was replaced or modified to form the ridged terrain. The almost identical crater frequencies exhibited by the cratered and ridged terrains indicate that they formed contemporaneously and over a relatively brief period of time. Graben formation was the next major event, beginning before the formation of the plains. Graben development may have continued even after plains formation had ended. As indicated by the range of crater frequencies, plains formation occurred over an extended period of time.

If the debris which cratered the uranian satellites was derived from heliocentric orbit or from orbits around Uranus that had a variety of orbital elements such that a given piece of debris could impact any of the satellites, then the cratering histories of the satellites can be correlated. If, however, the debris which cratered a given satellite was locally derived, say the destruction of co-orbiting satellites, then each satellite would have experien-

ced a unique cratering history and correlations would probably not be possible. If the former is the case¹, Ariel's history can be related to those of the other uranian satellites in at least a general way.

The observed crater frequencies for Ariel's cratered terrain are about one-tenth of those for either Oberon or Umbriel⁷, which indicates that Ariel's surface is significantly younger than the surfaces of either of these satellites. Relative to Titania⁸, Ariel's surface is less cratered, hence younger. Crater frequencies for Miranda⁶ are comparable to those of Ariel. Cratered terrain on Miranda has crater frequencies similar to the cratered and ridged terrains on Ariel. Arden and Elsinore Coronae on Miranda, the large circular regions of complex morphology and albedo on the leading and trailing hemispheres, have crater frequencies⁶ similar to those of the plains units on Ariel. Inverness Corona on Miranda has lower crater frequencies than any surface observed on Ariel. The similarity of crater frequencies suggests that geological activity on the two satellites was contemporaneous although Miranda's activity was longer lived.

If the crater-forming debris came from heliocentric orbits, then the cratering rate increase inwards from Oberon must be considered in the analysis of relative ages. Because of the focusing effect of the uranian gravitational field, the cratering rate at Ariel is about 5.4 times that at Oberon¹. The effect of the gradient would be to make the surface of Ariel younger relative to the surfaces of Umbriel, Titania or Oberon. Similarly, the cratering rate at Miranda is about 2.5 times that at Ariel, thus the effect would be to make Miranda's surface younger than Ariel's. In either case, Ariel was geologically active for longer periods of time than Oberon, Umbriel or Titania and for a similar length of time as Miranda.

The relation of crater frequencies to absolute time, the cratering rate, is quite speculative and model dependent. However, Smith *et al.*¹ have estimated the current cratering rate at Ariel to be about $4.3 \times 10^{-8} \text{ km}^{-2} \text{ yr}^{-1}$ for craters $\geq 10 \text{ km}$ in diameter. This cratering rate implies that Ariel's terrains formed in excess of 2.6 Gyr, probably near the end of the period of satellite accretion when the cratering rate was higher and exponentially decreasing. Thus, significant endogenically driven geological activity on Ariel would have occurred for a few hundred million years.

The tension indicated by the grabens suggests global expansion. Processes capable of producing global expansion are the freezing of a liquid interior of appropriate composition and global heating (obviously contrasting thermal regimes). Simple thermal expansion due to early heating should have a complimentary compressional episode due to subsequent cooling. Well-defined compressional features are not observed on Ariel, making a heating mechanism unlikely. Freezing of water or a eutectic melt of ammonia hydrate is a one-way process involving only the expansion that results from the volume increase due to freezing. Subsequent contraction due to continued cooling of the ice is negligible for a body of this size^{8,9} because no phase changes are encountered.

Ariel's thermal history, hence its geological history, may possibly be understood by a comparison with that of the saturnian satellite Dione, for which extensive thermal modelling has been done^{9,10}. Dione is a good analogue for Ariel because the two have a similar size, density and surface temperature. On the

basis of the thermal models of Dione, solid-state convection of the water-ice within Ariel's interior might be expected to last for several billion years. Temperatures in excess of the ammonia-water eutectic melting point (173 K) might develop near the surface of Ariel during the first few hundred million years after accretion and such temperatures might be sustained within the deep interior for up to 2 billion years. Methane-, argon-, and nitrogen-clathrates have dissociation points below 173 K, thus such materials could be mobilized for longer periods of time.

Stevenson and Lunine¹¹ have proposed a mechanism in which the viscosity of a fine-grained water-ice clathrate is reduced, perhaps to values of 10^{12} poise, by the presence of small amounts of intergranular cryogenic fluid. In comparison, water-ice at 90 K has a viscosity of $\sim 10^{33}$ poise. The reduced viscosity of such material would make it easier for it to reach the surface, particularly with the extensive network of faults that cuts the surface. Faults and fractures, especially deep-seated ones, can act as conduits through which molten material at depth can easily migrate to the surface.

Radioactive heating alone seems capable of producing a sufficient level of energy to drive most of the observed endogenic geological activity. It is not, however, capable of melting large volumes of water-ice, although it can generate large volumes of eutectic melts of ammonia hydrate. If graben formation reflects the freezing of water, then additional energy sources are necessary. Accretional heating^{13,14} may also have played an important role in the satellite's energy budget, particularly during accretion and immediately afterward. Energy deposited in the satellite during accretion may have promoted the initial destruction of the accretionary crust.

Tidal energy may also have added considerable amounts of energy to Ariel during its history. Although no tides occur at present, Ariel could have occupied several different resonances with Umbriel (ref. 14; S. J. Peale, personal communication) at some time in the past. Since Ariel is the more massive of the two, the energy would have been preferentially deposited in Ariel. However, most such resonances are stable and it becomes a problem to explain how Ariel escaped the resonance.

Summary

Crater size-frequency data and surface morphology indicate that Ariel has been completely resurfaced since its accretion. The cratered terrain, the oldest surface on Ariel, may have been the one formed during this initial global resurfacing. Subsequently, Ariel has been partly resurfaced as reflected by the ridged terrains and the plains. Similar crater frequencies for terrains other than the plain indicates that they are of similar age and formed over a relatively brief period of time. The plains are the youngest unit observed on Ariel. Graben development, reflecting an episode of global tension, began before plains formation, as indicated by the presence of plains material partly filling some grabens. A minimum age for graben formation is not well constrained by the observed stratigraphic relations. The absolute length of time during which endogenic activity occurred on Ariel is dependent on the cratering rate model, but probably lasted for only a few hundred million years.

The reviews and comments of J. Veverka as well as discussions with G. Shoemaker, L. Soderblom and J. Boyce are greatly appreciated.

Received 29 December 1986; accepted 3 April 1987.

- Smith, B. A. *et al. Science* **233**, 43–64 (1986).
- Brown, R. H. in *Uranus and Neptune* NASA Conf. Publ. 2330 (ed. Bergstrahl, J. T.) 437–461 (NASA, Washington DC, 1984).
- Brown, R. H. & Cruikshank, D. in *Satellites* (eds Burns, J. & Matthews, M. S.) (University of Arizona Press, Tucson, in the press).
- Smith, B. A. *et al. Science* **215**, 504–537 (1982).
- Shoemaker, E. M. *et al. in Satellite of Jupiter* (ed. Morrison, D.) 435–520 (University of Arizona Press, Tucson, 1982).
- Plescia, J. B. in *Abstr. Eighteenth Lunar and Planetary Science Conf.* 784–785 (Lunar and Planetary Institute, Houston, Texas, 1987).

- Plescia, J. B. in *Abstr. Eighteenth Lunar and Planetary Science Conf.* 786–787 (Lunar and Planetary Institute, Houston, Texas, 1987).
- Consolmagno, C. J. *Icarus* **64**, 401–443 (1985).
- Ellsworth, K. & Schubert, G. *Icarus* **54**, 490–510 (1983).
- Stevenson, D. J. *Nature* **298**, 142–144 (1982).
- Stevenson, D. J. & Lunine, J. I. *Nature* **323**, 46–48 (1986).
- Squyres, S. W., Reynolds, R. T. & Lissauer, J. J. *Icarus* **61**, 218–223 (1985).
- Stevenson, D. J. in *Uranus and Neptune* NASA Conf. Publ. 2330 (ed. Bergstrahl, J. T.) 405–424 (NASA, Washington DC, 1984).
- Squyres, S. W., Reynolds, R. T., Summers, A. L. & Shung, F. in *Abstr. Eighteenth Lunar and Planetary Science Conf.* 944–945 (Lunar and Planetary Institute, Houston, Texas, 1987).

New partial skeleton of *Homo habilis* from Olduvai Gorge, Tanzania

Donald C. Johanson*, Fidelis T. Masao†, Gerald G. Eck‡, Tim D. White§, Robert C. Walter||, William H. Kimbel*, Berhane Asfaw§, Paul Manega¶, Prosper Ndessokia* & Gen Suwa§

* Institute of Human Origins, 2453 Ridge Road, Berkeley, California 94709, USA

† National Museums of Tanzania, Box 511, Dar es Salaam, Tanzania

‡ Department of Anthropology, University of Washington, Seattle, Washington 98105, USA

§ Department of Anthropology, University of California, Berkeley, California 94720, USA

|| Center for Geochronological Research, INSTAAR, Box 450, University of Colorado, Boulder, Colorado 80309, USA

¶ Department of Geology, University of Colorado, Boulder, Colorado 80309, USA

Department of Paleontology, University of California, Berkeley, California 94720, USA

A new partial skeleton of an adult hominid from lower Bed I (about 1.8 Myr ago), Olduvai Gorge, is described. This specimen's craniodental anatomy indicates attribution to Homo habilis, but its postcranial anatomy, including small body size and relatively long arms, is strikingly similar to that of some early Australopithecus individuals.

OLDUVAI Gorge in northern Tanzania has yielded important evidence bearing on human origins and evolution. Discoveries made from the 1950s to the 1970s have demonstrated the coexistence of two hominid taxa in late Pliocene and early Pleistocene strata at the site: *Australopithecus boisei* and *Homo habilis*¹⁻³. Palaeoanthropological research at Olduvai has been carried out since 1985 by a team from the National Museums of Tanzania, the Tanzanian Department of Antiquities, the Institute of Human Origins and the University of California, Berkeley. We report here the discovery of a new hominid specimen from lower Bed I at Olduvai Gorge.

Recovery

On 21 July 1986, the third day of surface survey of the 1986 field season, one of us (T.D.W.) discovered a fragment of hominid ulna on the surface of Bed I sediments at FLK, near Geological Locality 45c (ref. 4) and 25 m west of the road to FLK Zinj (Fig. 1a and b). A search and screening of the adjacent surface resulted in the recovery of maxillary, calvarial, mandibular, radial, humeral, femoral and tibial fragments of what appears to be a single hominid individual. This specimen is designated Olduvai Hominid (OH) 62.

Stratigraphy and geological framework

The site at which OH 62 was found lies roughly 250 m south-east of the FLK (*Zinjanthropus*) locality (Fig. 1a). Hominid fragments were found on the surface and in a thin colluvial soil (Fig. 1b and c), over an area of about 40 m² on the north side of a small knoll, Dik Dik Hill (DDH).

Roughly 16 m of deposits are exposed in the gorge wall immediately south-west of the hominid discovery. These consist of Bed I and lower Bed II sediments unconformably capped by the Ndutu Beds. Lower Bed I lava forms the present-day valley floor. Roughly 2 m of beige tuffaceous silts and clays, which are subjacent to Tuffs IC and ID, overlie the lava.

The base of the section at the hominid site (Trench 2) consists of thinly laminated beige and brown clays (Unit 2-1), which are overlain by 30 cm of tuffaceous silts and silty clays containing large, altered pumice clasts (up to 3 cm) in the basal part (Unit 2-2) (Fig. 1b and c). This sequence is unconformably capped by a 0-50 cm colluvium containing large lithic clasts. All of the vertebrate (including hominid) fossils were on top of or within this colluvium.

The colluvium capping DDH (Unit 2-3, Fig. 1c) probably formed as a lag deposit after a section of the gorge wall collapsed and deflated. Tuff ID forms a resistant ledge in the local topogra-

phy, and there are numerous examples nearby where portions of this ledge have calved away from the gorge wall, leaving a lag similar to that observed at DDH. More than 95% of the lithic clasts in the colluvium (Trench 2) are positively correlated with deposits equivalent to and below Tuff ID (Trench 1).

The occurrence of OH 62 in colluvial soil on the flank of DDH and the surface distribution pattern of numerous associated fragments of one hominid individual indicate that OH 62 eroded from DDH. In addition, *in situ* nonhominid fossils were found in a dark, discontinuous laminated sand immediately below Tuff IC in Trench 3 (Fig. 1b). These remains are similar in preservation to OH 62. Thus, through several indirect yet highly congruent lines of evidence we conclude that OH 62 derives from lower Bed I deposits below Tuff ID, most probably from the sand lens below Tuff IC, roughly equivalent to the FLK (*Zinjanthropus*) level⁴. This implies an age for OH 62 of between 1.85 Myr, the age of Tuff IB (ref. 5; using revised K—Ar constants, ref. 6), and 1.75 Myr, the inferred age of Tuff IF⁴. There are no firm isotopic ages for Tuffs IC or ID, although tuffs between ID and IC are dated to 1.80 Myr (± 0.08 , 1 S.D.)⁵.

The hominid bones, including entire shafts, shattered as they eroded from the deposit and deflated as particles in a colluvial lag, or desert pavement. We estimate on geomorphological grounds that the process of exposure took centuries.

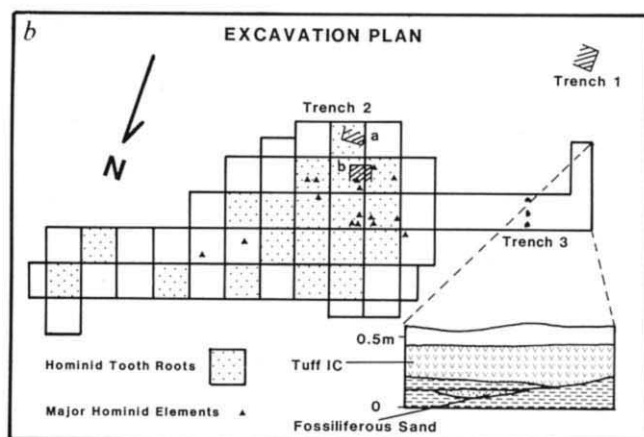
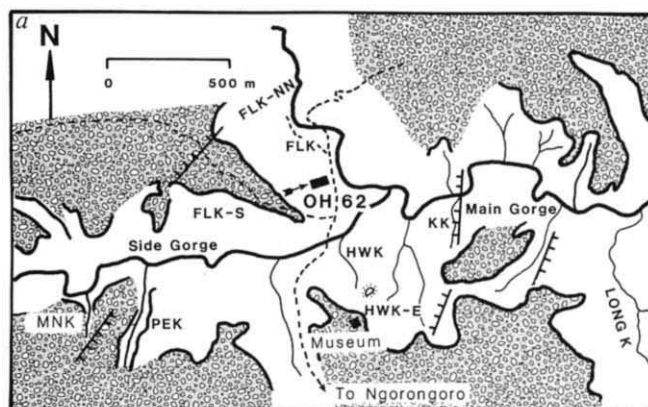
Associations

Screening and excavation recovered nearly 18,000 fragments of fossil bone and tooth. The hominid fossils are highly lithified and dark grey to black in colour, distinguishing them from specimens derived from higher in the section. Most of the nonhominid fossils associated with OH 62 are small fragments of small to medium sized mammals and reptiles. These include: *Kobus sigmoidalis*, cf. Antelopini, cf. Alcelaphini, Giraffidae (cf. *Sivatherium*), Hippopotamidae, Suidae, Deinotheriidae, Carnivora (cf. large lutrine mustelid or viverrid), Cercopithecidae, Hystricidae, Cricetidae, Aves (cf. *Struthio*), Reptilia (Varanidae, *Crocodylus*, Serpentes, Chelonia), Amphibia (cf. Anura) and Pisces (*Clarius*).

Artefacts recovered on the surface and in the lag deposit are cores and flakes typical of the Oldowan industry, but whether they are contemporary with the hominid cannot be established. No stone tools were found in the outcrop of the laminated sand unit.

Preliminary description of the skeleton

After full conjoining of the recovered pieces it is evident that



c MICROSTRATIGRAPHY OF THE HOMINID SITE

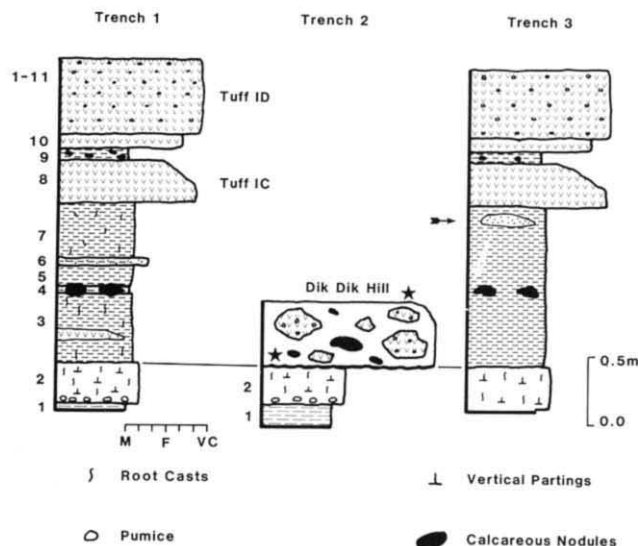


Fig. 1. *a*, Plan view of the junction of the Main and Side gorges, showing the location of the OH 62 hominid site. Patterned areas represent high grassy plains. Gorge sediments are depicted in white, adjacent to stream channels. Several major faults are shown, with hatchure marks indicating down-throw directions. Dashed lines represent roads. *b*, Excavation plan for OH 62 site, showing distribution pattern of hominid remains within the 2 m × 2 m grid system. Geological trenches discussed in the text are illustrated. The summit of Dik Dik Hill occurs between Trench 2a and 2b. The westernmost edge of Trench 3 exposes a 0.5 m section of Tuff IC that is underlain by a dark discontinuous channel sand (inset). This sand contains vertebrate fossils with preservation features like those of OH 62. *c*, Detailed stratigraphy of the hominid site (Dik Dik Hill) and adjacent sediments, as discussed in the text. The hominid level is indicated by the stars, and the fossiliferous channel sand by the arrow. The lateral dimension of each unit refers to clast size according to the scale at the base of the diagram: mud (M), fine (F), very coarse (VC).

parts of the skull, right arm and both legs are represented (Table 1). There is no duplication of elements, which suggests the presence of only one individual. Preservation and colour of all fragments correspond and all elements are fully adult.

All limb bone articular ends are missing except for a portion of the proximal ulna. The bones show no evidence of rodent- or carnivore-induced damage and display no cutmarks or

evidence of peri-mortem fracture. The bone surfaces are slightly weathered and suggest limited exposure before deposition.

Skull. The skull of OH 62 is represented by portions of the palate, face, calvaria, mandible and dentition. The following descriptions focus only on features of taxonomic importance. *Palate* (Fig. 2a). The palatal surface is preserved between the incisor alveoli and the M2/M3 level, but the alveolar processes

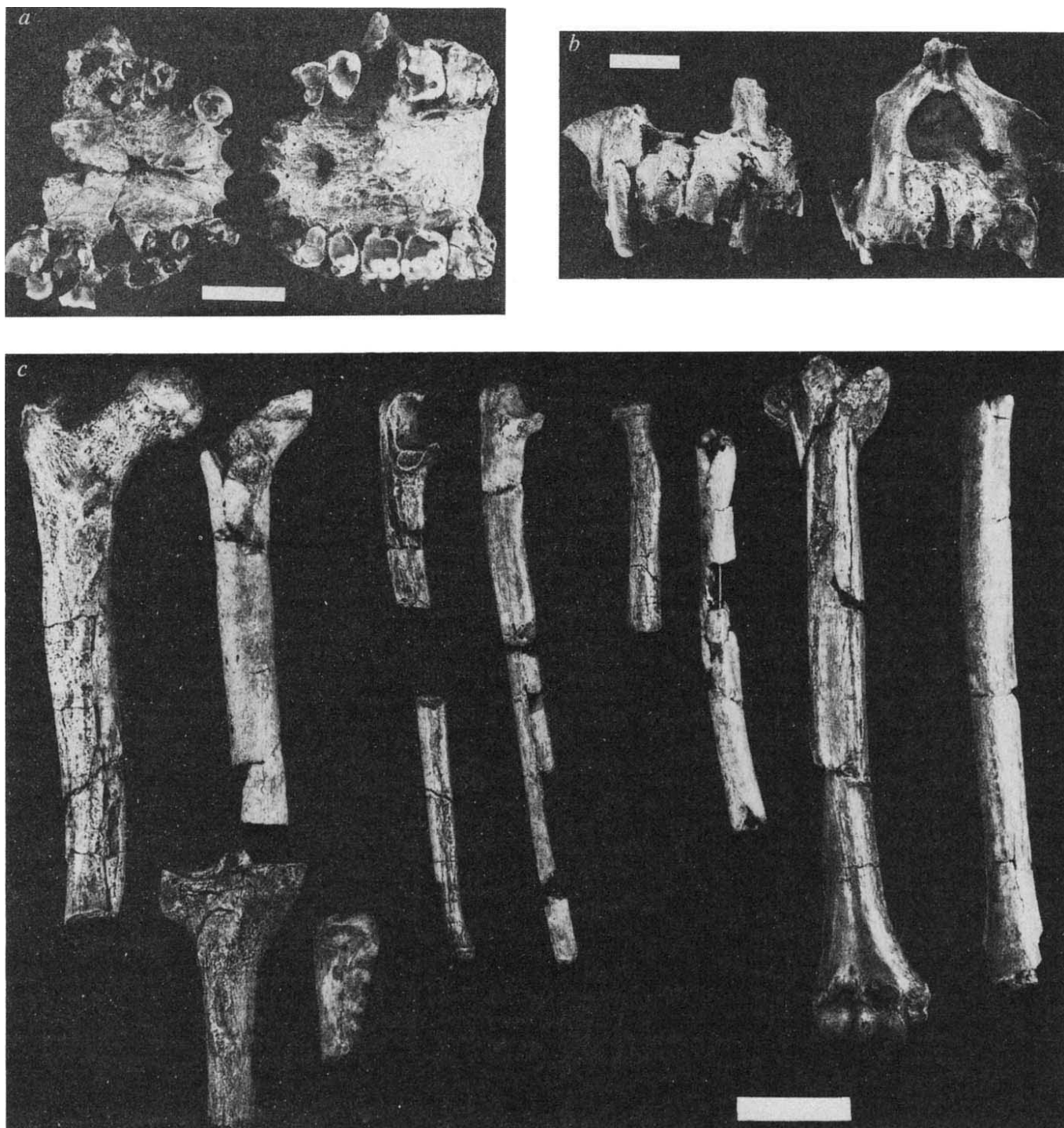


Fig. 2 *a*, Palatal views of OH 62 (left) and Stw. 53 (cast, right) maxillae. Bar, 2 cm. *b*, Facial views of OH 62 (left) and Stw. 53 (cast, right). Bar, 2 cm. *c*, Postcranial elements of OH 62 compared with those of *A. afarensis* specimen A.L. 288-1 ('Lucy'). The Olduvai fossils are to the right in each comparison. Bar, 4 cm.

are broken at most tooth positions. The palate is moderately deep posteriorly, and shelves inferiorly anterior to the large incisive fossa. As judged by estimates of internal breadth (~35 mm) and length (~41 mm, measured to mid-M² level), it is clear that the palate is relatively wide compared to *Australopithecus* (for example, A.L. 200-1a, Sts. 5, Sts. 52a, OH 5). In size and morphology, the OH 62 palate is similar to OH 24, and especially to Stw. 53, a *Homo habilis* skull from the Sterkfontein (Member 5) Extension Site breccia in South Africa^{7,8}. **Face** (Fig. 2*b*). Most of the region around the nasal cavity, including the right zygomatic process of the maxilla, is pres-

ved. The maxilla is moderately prognathic. It is evident that the nasoalveolar clivus is flat, short and minimally projecting relative to the bicanine line. The eroded remnant of a small, but distinct, anterior nasal spine is present at the entrance to the nasal cavity. The spine is separated from the nasal orifice of the incisive canals, and hence from the inferred anterior insertion of the vomer, by a 6.0 mm long, horizontal intranasal platform. Laterally, the inferior nasal margin demarcates a fairly distinct change in contour between the clivus and the nasal cavity floor.

The anterolaterally facing nasal process of the maxilla shows that superiorly the lateral margin of the nasal aperture was sharp

and everted. An 'anterior pillar' is not present, and there are no indications of a distinct canine fossa or 'maxillary furrow' intervening between the nasal and zygomatic processes. The zygomatic process arises low on the maxilla. Its root is positioned posteriorly, at the M1 level.

The isolated frontal process of the left zygomatic is slender and bears a weak postmarginal process on its temporal face. The facial plate of the process, separated from the orbital plate by a sharp lateral orbital margin, faces distinctly laterally.

In almost all these characters, the OH 62 facial skeleton closely resembles specimens attributed to *Homo habilis* or *Homo sp.* (for example OH 24, KNM-ER 1470, 1813, SK 847, and especially Stw. 53). Specializations defining *Australopithecus africanus*, *A. robustus* and *A. boisei* are notably absent in the new Olduvai face^{9,10}.

Calvaria. Very little of the cranial vault of OH 62 survives. Preserved fragments of the sphenoid, occipital, and other vault bones reveal no taxonomically valuable morphology.

Mandible. The preserved right condylar neck is gracile and the condyle is small compared to 'robust' *Australopithecus* specimens. A basal fragment evinces a degree of eversion similar to that seen in *H. habilis* mandible KNM-ER 1802, although the OH 62 mandible base was obviously less robust.

Dentition. Molar wear pattern and anterior/posterior dental proportions (as estimated from maxillary alveoli and preserved canine and molar crowns) indicate that OH 62 does not represent a 'robust' *Australopithecus*. Both P³ and P⁴ possess double buccal roots. These roots lack the massiveness of 'robust' *Australopithecus* homologues and are similar to the condition seen in Stw. 53.

The right C measures 9.9 mm buccolingually. The buccolingual breadths of the left M² and M³ are estimated to be ~16.0 mm and ~15.5 mm, respectively. These molar dimensions approximate those of OH 16 and Stw. 53, and lie in the upper range of early *Homo* specimens. The canine/molar breadth ratio is therefore at the low end of the *Homo* range.

Other morphological features, such as a virtual lack of asymmetry in upper canine occlusal outline, a simple distal buccal groove on the lower canine lacking a 'V' configuration, and M³ with a relatively short mesiodistal dimension lingually, suggest that OH 62 represents *Homo* rather than *A. africanus*.

Postcranium. Description of the OH 62 postcranium will focus on comparisons with the A.L. 288-1 partial skeleton from Hadar ('Lucy'; ref. 11) because of the size and anatomical similarities between these two hominid individuals (Fig. 2c).

Humerus. The OH 62 humeral shaft is essentially intact, with most of the bicipital groove present. The proximal shaft circumference closely approximates that recorded for A.L. 288-1, whereas the distal shaft of OH 62 is slightly thinner. We conservatively estimate the OH 62 humeral length at ~27 mm longer than that of A.L. 288-1. The total OH 62 upper arm was almost certainly longer than that of 'Lucy'.

Radius. Most of the radial tuberosity is intact and about two-thirds of the shaft below this point is preserved. The shaft exhibits moderate mediolateral bowing and a weak interosseous crest. The radial tuberosity is larger, more rounded and less divided than that of the A.L. 288-1 radius. Shaft circumference is greater than that measured for A.L. 288-1 and the radius of OH 62 is slightly more robust in comparable parts.

Ulna. The proximal ulna segment retains the inferior extension of the trochlear notch and a bit of eroded radial notch. The entire anatomy of the OH 62 proximal ulna is very similar to that recorded for A.L. 288-1. Mediobasal and anteroposterior measures of OH 62 and A.L. 288-1 at the level of the radial notch differ by less than 1 mm and shaft circumferences at the base of the brachialis insertion are the same. The distal OH 62 shaft fragment shows part of a pronator quadratus insertion. Shaft circumference at the distal end of this line (minimum shaft circumference) is 2 mm less than that measured on A.L. 288-1.

Femur. A visible obturator externus groove marks the posterior

Table 1 List of remains attributed to OH 62

A.	Partial maxilla with RC, partial RP ⁴ , RM ¹ mesiobuccal root fragment; partial LC, LP ³ root fragment, LP ⁴ root fragment, partial LM ² -LM ³ (32 conjoined pieces)
B.	Frontal process of left zygomatic
C.	Posterior wall fragment of zygomatic process, right maxilla
D.	Posterior wall fragment of zygomatic process, left maxilla
E.	Maxillary and palatine fragments (15 pieces)
F.	Fragment of greater wing of right sphenoid
G.	Greater wing fragment of left sphenoid
H.	Occipital fragment with superior sagittal sulcus
I.	Cranial vault fragments (21 pieces)
J.	Right mandibular condyle
K.	Anteromedial wall of left mandibular corpus
L.	Posteromedial fragment of left mandibular corpus
M.	Basal fragment of right mandibular corpus
N.	Mandible fragments (17 pieces)
O.	Root and distal crown fragment of RC
P.	Root of LC
Q.	Root of LP ₄
R.	Lingual root of RP ³
S.	Buccal fragment of RM ¹ crown
T.	Labial fragment of I ¹ (side indeterminate)
U.	Dental crown and root fragments (position and side indeterminate) (161 pieces)
V.	Right humerus shaft (5 conjoined pieces)
W.	Right radius shaft (13 conjoined pieces)
X.	Proximal, midshaft, and distal fragments of right ulna (4, 6 and 1 conjoined pieces)
Y.	Proximal shaft and neck of left femur (10 conjoined pieces)
Z.	Proximal fragment of right tibia
Total number of pieces = 302	

surface of the femur neck. The estimated neck/shaft angle of OH 62 is roughly the same as that documented for A.L. 288-1 (123°). The inferior half of the neck suggests the neck was flattened anteroposteriorly and was at least as long as that of A.L. 288-1. Anteroposterior and mediolateral shaft diameters at the base of the OH 62 lesser trochanter are both 21 mm, whereas homologous measures for the more anteroposteriorly compressed A.L. 288-1 shaft are 27 mm and 18 mm, respectively. Shaft circumference of the OH 62 specimen near midshaft is 9 mm less than that measured on A.L. 288-1. Even allowing for slight exfoliation of the OH 62 femur, visual comparison makes it obvious that this individual's femur was smaller and less robust than the A.L. 288-1 femur.

Tibia. Only the tibial tuberosity and the shaft immediately distal to it are preserved. The tuberosity is large and the surface lateral to it is deeply excavated as in A.L. 288-1. The tuberosity is well demarcated from the intracapsular area above by a marked transverse groove similar to that seen on the A.L. 288-1 and A.L. 129-1b proximal tibiae. Unlike these and other Hadar tibiae, however, OH 62 lacks any pit associated with the tibialis anterior muscle.

Discussion

The new OH 62 partial skeleton from Bed I represents a significant addition to the hominid fossil record because of its bearing on several key issues concerning *Homo habilis*, a taxon originally created on the basis of fragmentary cranial and questionably associated postcranial elements from Olduvai Gorge².

Several important conclusions about the OH 62 partial skeleton emerge from the first round of analysis. First, body size for this fully adult individual is estimated to be as small as or smaller than that of any known fossil hominid. Second, in addition to size, there are striking anatomical and proportional similarities between the OH 62 postcranial skeleton and small *Australopithecus* individuals (especially A.L. 288-1). Third, the strong morphological similarities of the OH 62 face, palate and dentition to *Homo habilis* (especially Stw. 53) warrant attribution of the Olduvai individual to this taxon. This represents the

first time that limb elements have been securely assigned to *Homo habilis*.

Taxonomy. The original description of Bed I *Homo habilis* elicited assertions that species-level distinction from *Australopithecus africanus* was unwarranted¹²⁻¹⁴. More complete cranial material from Koobi Fora (KNM-ER 1470, 1590, 1813, 3732) and Sterkfontein (Stw. 53) has shown that *Homo habilis* does indeed merit specific distinction from *A. africanus*^{7,10,15}. However, this more complete material has led some investigators to divide the larger collection into *A. africanus* (such as KNM-ER 1813, OH 13), and *Homo* (such as KNM-ER 1470, 1590) sub-samples¹⁶. Others have suggested a similar division but discern at least two "non-australopithecine" species in the sample^{17,18}.

Many workers continue to follow Tobias¹⁹ in placing all of the late Pliocene-early Pleistocene, non-'robust' specimens that fall outside the range of *Homo erectus* into a variable, polymorphic species *Homo habilis*²⁰⁻²⁴. This view holds that morphological differences between individuals such as KNM-ER 1470 and 1813 reflect sexual, geographical and chronological factors within a single species. Given the evidence currently available, we agree and therefore view OH 62 as an elderly member of *Homo habilis*. The similarity between comparable Stw. 53 and OH 62 cranial and dental parts shows that this taxon was widespread in Africa.

Differentiation of *Australopithecus* and early *Homo* species at sites such as Olduvai, Omo and Koobi Fora has been accomplished exclusively on the basis of cranial, mandibular and dental characters. Taxonomic attribution of the wide variety of isolated postcranial elements found at these sites has been made only through speculation. Pilbeam's recent review reveals this when he states: "There are no clear associations between skulls of *Homo habilis* and the other bones of the species, but the limb bones that are assumed to represent it, unlike those of *Australopithecus*, resemble those of species of the genus *Homo* (with the exception of modern *Homo sapiens*)"²⁵.

On the contrary, the new OH 62 partial skeleton as well as the juvenile hand parts described as part of the holotype (OH 7; ref. 26) show that *Homo habilis* had a postcranium similar in many ways to mid-Pliocene *A. afarensis*. Until now, late Pliocene-early Pleistocene postcrania have been sorted into *Homo* or 'robust' *Australopithecus* species on the basis of size and assumed limb proportions. This practice is seriously undermined by OH 62. Many previous taxonomic attributions based on postcrania should thus be regarded as highly tenuous.

Functional anatomy. Susman and Stern's assessment of the limited postcranial remains of *Homo habilis* led them to characterize members of this taxon as habitual bipeds who retained primitive characters in the hand but were "... less arboreal and more modern in their bipedality than the hominids of 3-4 million years ago."²⁷ Allegedly primitive features of the Olduvai *H. habilis* foot (OH 8) have been noted by Wood²⁸ and Lewis²⁹. The discovery of long and powerful arms in *Homo habilis* and *Australopithecus*-like aspects of pelvic and proximal femoral

anatomy in early *Homo erectus* (KNM-WT 15000; ref. 30) emphasize the mosaic pattern of evolution in the early hominid postcranial skeleton. However, the juxtaposition of an otherwise relatively derived *H. erectus* postcranium at ~1.6 Myr (KNM-WT 15000) and a postcranially primitive *H. habilis* at ~1.8 Myr (OH 62) may imply an abrupt transition between these taxa in eastern Africa.

We estimate the OH 62 humerus length at 264 mm, 27 mm longer than the A.L. 288-1 humerus¹¹. If the length of the less robust OH 62 femur was no greater than that of A.L. 288-1 (280 mm), the Olduvai individual would have a humerofemoral index of close to 95%. Further work on this intriguing issue is in progress.

The new partial skeleton OH 62 offers additional evidence regarding the relative size of the postcanine dentition. McHenry has concluded that postcanine megadontia (relative to body size estimates) was reduced in *Homo habilis* compared to the *Australopithecus* condition³¹. Our measurements indicate that the length of the postcanine tooth row relative to body size in OH 62 was at least as great as that of the comparably small A.L. 288-1 specimen of *Australopithecus afarensis*. Thus, the degree of megadontia in *Homo habilis* may, in fact, be little changed from the *Australopithecus* condition.

The very small body size of the OH 62 individual suggests that views of human evolution positing incremental body size increase through time may be rooted in gradualistic preconceptions rather than fact³². It is not possible to estimate cranial capacity for OH 62, but this skeleton and other available data on cranial capacity suggest the possibility that small individuals of *Australopithecus* and *Homo habilis* were differentiated by cranial capacity but not by body size. This reinforces the view that encephalization in the terminal Pliocene played a key role in hominid evolution^{19,33}.

We thank the Tanzanian Government, including the National Scientific Research Council, the National Museums, the Department of Antiquities, and the Ngorongoro Conservation Area for encouragement, support and permission to conduct research at Olduvai and Laetoli. This new discovery emphasizes the need for continuing management of Tanzanian sites such as Olduvai Gorge. Special thanks are due to Professor Msangi and Professor Hirji, Dr Waane, and Mr Mturi. Support of the faculty and associates of the University of Dar es Salaam, including J. Karoma and P. Schmidt of the Archaeology Unit and Professor Aswathanaratana, and others is appreciated. Advice and field assistance was provided by Lew and Nancy Binford, and George and June Frison. Field support and assistance was provided by Sandy Evans of Abercrombie and Kent, E.A. von Mutius, Peter Lauwo, Lucas Tarimo, Stanley Nderingo, James Shreeve, Jeremy Paul, Mzee Mrisho, Alberto Angela, and Hans Schneider. Thanks go to Owen Lovejoy for assistance with the postcranium, and to Mary Leakey for providing the specimen number for the new hominid. We are especially grateful to Gordon Getty, David Koch, and Ligabue Research and Study Center for their financial support.

Received 3 March, accepted 24 March 1987.

1. Leakey, L. S. B. *Nature* **184**, 491-493 (1959).
2. Leakey, L. S. B., Tobias, P. V. & Napier, J. R. *Nature* **202**, 7-9 (1964).
3. Tobias, P. V. *The Cranium and Maxillary Dentition of Australopithecus* (Zinjanthropus) boisei, *Olduvai Gorge Volume 2* (Cambridge University Press, 1967).
4. Hay, R. L. *Geology of the Olduvai Gorge* (University of California Press, Berkeley, 1976).
5. Curtis, G. H. & Hay, R. L. in *Calibration of Hominoid Evolution* (eds Bishop, W. W. & Miller, J. A.) 289-301 (Scottish Academic, Edinburgh, 1972).
6. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359-362 (1977).
7. Hughes, A. R. & Tobias, P. V. *Nature* **256**, 310-312 (1977).
8. Tobias, P. V. *Z. Morph. Anthropol.* **69**, 225-265 (1978).
9. Clarke, R. J. thesis, Univ. Witwatersrand (1977).
10. Rak, Y. *The Australopithecine Face* (Academic, New York, 1983).
11. Johanson, D. C. et al. *Am. J. Phys. Anthropol.* **57**, 403-452 (1982).
12. Robinson, J. T. *Nature* **205**, 121-124 (1965).
13. Robinson, J. T. in *Evolutionary Biology Vol. 1* (eds Dobzhansky, T., Hecht, M. K. & Steere, W. C.) 69-100 (Appleton-Century-Crofts, New York, 1967).
14. Brace, C. L., Mahler, P. E. & Rosen, R. B. *Yearb. Phys. Anthropol.* **16**, 50-68 (1973).
15. Wood, B. A. in *Recent Advances in Primatology Vol. 3* (eds Chivers, D. J. & Joysey, K. A.) 351-372 (Academic, London, 1978).

16. Walker, A. C. & Leakey, R. E. *Sci. Am.* **239**, 54-66 (1978).
17. Wood, B. A. in *Ancestors: The Hard Evidence* (ed. Delson, E.) 206-214 (Liss, New York, 1985).
18. Stringer, C. B. in *Major Topics in Primate and Human Evolution* (eds Wood, B. A., Martin, L. & Andrews, P.) 266-294 (Cambridge University Press, 1986).
19. Tobias, P. V. *Can. J. Anthropol.* **3**, 163-185 (1983).
20. Johanson, D. C. & White, T. D. *Science* **203**, 321-330 (1979).
21. Kimbel, W. H., White, T. D. & Johanson, D. C. *Am. J. Phys. Anthropol.* **64**, 337-388 (1984).
22. Skelton, R. R., McHenry, H. M. & Drawhorn, G. M. *Curr. Anthropol.* **27**, 21-43 (1986).
23. Howell, F. C. in *Evolution of African Mammals* (eds Maglio, V. M. & Cooke, H. B. S.) 154-258 (Harvard University Press, Cambridge, 1978).
24. Wolpoff, M. H. *Paleoanthropology* (Knopf, New York, 1980).
25. Pilbeam, D. R. *Sci. Am.* **250**, 84-97 (1984).
26. Susman, R. L. & Stern, J. T. *Science* **217**, 931-934 (1982).
27. Stern, J. T. & Susman, R. L. *Am. J. Phys. Anthropol.* **60**, 279-317 (1983).
28. Wood, B. A. *J. hum. Evol.* **3**, 373-378 (1974).
29. Lewis, O. J. *J. Anat.* **131**, 275-298 (1980).
30. Brown, F., Harris, J., Leakey, R. & Walker, A. *Nature* **316**, 788-792 (1985).
31. McHenry, H. M. *Am. J. Phys. Anthropol.* **64**, 297-306 (1984).
32. Cronin, J. E., Boaz, N. T., Stringer, C. B. & Rak, Y. *Nature* **292**, 113-122 (1981).
33. Holloway, R. L. *Hum. Neurobiol.* **2**, 105-114 (1983).

Origin of the large-scale galaxy peculiar velocity field: a minimal isocurvature model

P. J. E. Peebles

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, USA

The galaxy peculiar velocity field (velocity relative to the uniform Hubble law expansion) seems to have a large r.m.s. value, $\sim 500 \text{ km s}^{-1}$, and a large coherence length $\geq 50 \text{ Mpc}$. It is difficult to account for this in the standard model inspired by inflation, where galaxies and clusters of galaxies grow by gravity out of initially adiabatic fluctuations in the distribution of mass now dominated by weakly interacting cold dark matter^{1,2}. A model that does seem to work assumes that, at high redshift, mass was uniform and the baryon number was inhomogeneous, a stationary random process. This produces in the late Universe a baryon distribution with a power spectrum proportional to the initial value at short wavelength, a prominent peak in the spectrum at the matter-radiation Jeans length, and a sharp cutoff at longer wavelengths. Here I report how this spectrum normalized by the observed galaxy clustering can produce by gravity the large velocity field with large coherence length needed without excessively perturbing the microwave background.

A velocity autocorrelation function is

$$v_c(r)^2 = \langle \mathbf{v}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x}+\mathbf{r}) \rangle \quad (1)$$

where $\mathbf{v}(\mathbf{r})$ is the galaxy peculiar velocity field. It is thought that $v_c(0) \sim 500 \text{ km s}^{-1}$ and $v_c(l) = 0.5v_c(0)$ at coherence length $l \sim 50\text{--}100 \text{ Mpc}$. This would account for the fact that the Local Group and all our neighbours at cosmological redshifts $cz \leq 800 \text{ km s}^{-1}$ have a roughly common motion $\sim 600 \text{ km s}^{-1}$ relative to the microwave background^{3,4} and for the large mean peculiar motions seen in the Rubin-Ford^{5,6} and Burstein *et al.*⁷ samples. As discussed by Vittorio *et al.*¹, and Bond², it is difficult to account for the large value of l if large-scale mass fluctuations grow out of initially adiabatic fluctuations with the scale-invariant spectrum, P , proportional to wave number k . Tilting the spectrum to $P \propto k^n$, $n \leq 0$, would increase l , but would also suppress the amplitude of density fluctuations on the scale of galaxies, which would delay galaxy formation, which already occurs at dangerously low redshifts if $n = 1$ (ref. 8).

This impasse (among other considerations⁹) suggests that an alternative possibility should be examined, namely a minimal isocurvature model. This model assumes mass is dominated by the usual baryons, radiation and three massless neutrinos each with two spin states, which is minimal in that it ignores non-baryonic, non-relativistic matter. The second assumption is that in the very early Universe the mass is homogeneous, so space curvature is unperturbed, but the baryon number is clustered. This is not motivated by the standard particle physics picture of the early Universe; it is introduced in an attempt to match the observations.

Evolution tends to preserve the initially homogeneous mass distribution in this model, so when the Universe has become matter-dominated the baryon distribution is close to homogeneous on scales large compared to the matter-radiation Jeans length. On small scales radiation pressure resists this so that when matter and radiation decouple the baryons end up with their original distribution. The result is a power spectrum proportional to the original on small scales and cutoff at large scales¹⁰. The advantage of the cutoff is that it can be assumed that the spectrum increases with increasing wavelength to enhance mass fluctuations on the scale of superclusters, without having unacceptably large mass fluctuations on the scale of the horizon.

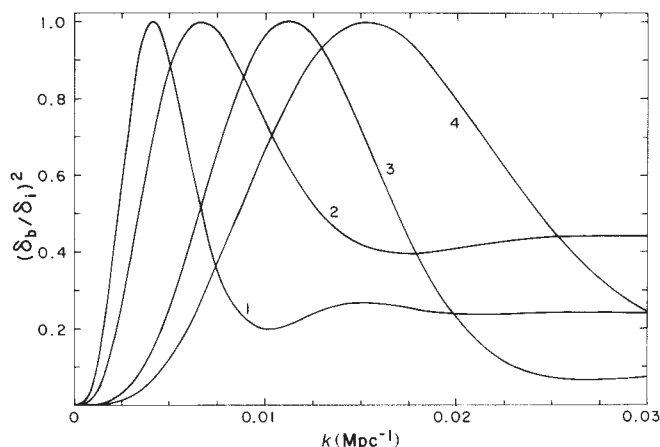


Fig. 1 Baryon power spectrum transfer function. The vertical axis is the ratio of final power spectrum at $y \gg 1$ to the initial spectrum, in linear perturbation approximation, and plotted on an arbitrary scale. The horizontal axis is the wave number in radians per Mpc at the present epoch. Curves 1 and 3 assume that early star formation keeps the matter fully ionized. Curves 2 and 4 assume that the fractional ionization $x(t)$ is the larger of thermal equilibrium and $x_{\min} = 0.1$. For curves 1 and 2, $\Omega h^2 = 0.025$; for 3 and 4, $\Omega h^2 = 0.1$.

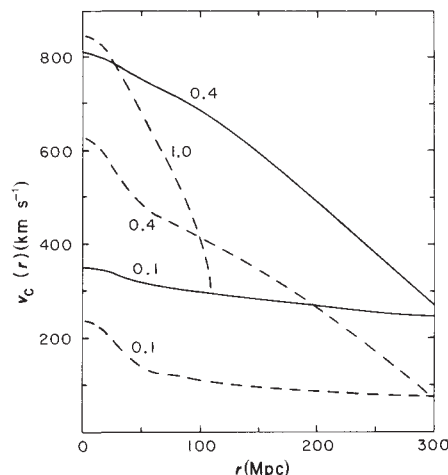


Fig. 2 Velocity correlation function (equation (1)). The label is the density parameter, with $h = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The primeval spectrum of the baryon distribution is $P_i \propto k^n$, with $n = 0$ (dashed lines) or $n = -1$ (solid lines).

The equations describing the evolution of the distributions of matter and radiation are written down in ref. 11. At very long wavelengths the density contrast in neutrinos and radiation is $\delta_r = \frac{4}{3}(\delta_b - \delta_i)$, where δ_b is the baryon density contrast and $\delta_b(t) = \delta_i$ at very high redshift. The solution for $\delta_b(t)$ is then

$$\frac{\delta_b}{\delta_i} = \frac{4}{y} - \frac{8}{y^2} [(1+y)^{1/2} - 1] \quad \text{where } y = \frac{\rho_b(t)}{\rho_r(t)} \quad (2)$$

where ρ_b and ρ_r are the mean mass densities in baryons and relativistic matter. At shorter wavelengths the equations can be integrated numerically¹¹ with initial conditions from equation (2). There is one free function, the ionization history $x(t)$, and one parameter, the product Ωh^2 , where Ω is the density parameter (ratio of present mass density to Einstein-de Sitter density) and h is Hubble's constant in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The result shown in Fig. 1 is the power spectrum transfer func-

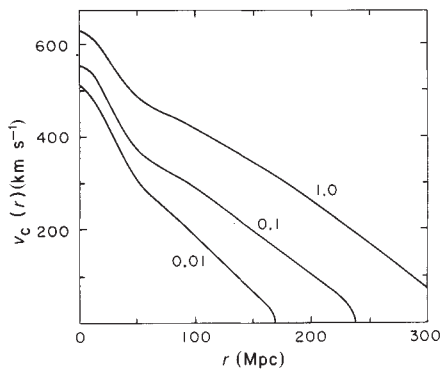


Fig. 3 Effect of ionization history on the velocity correlation function for the case $\Omega = 0.4$, $n = 0$. The parameter is the minimum ionization.

tion, the ratio $(\delta_b/\delta_i)^2$, where δ_i is the initial value and δ_b is the value at $y \gg 1$ and after matter has decoupled from radiation.

The prominent peak in the transfer function appears at about the matter-radiation Jeans length, where the radiation density contrast $\delta_r(t)$ completes about one half oscillation before the baryons decouple from the radiation. This partial oscillation makes $d\delta_b/dt$ large at decoupling, which boosts the growth of $\delta_b(t)$. The position of this peak scales as $k_{\max} \sim \Omega h^2$. At $k > k_{\max}$ the oscillation of $\delta_r(t)$ is strongly damped, so at decoupling $\delta_b \approx \delta_i$ and $d\delta_b/dt \approx 0$, so the transfer function is very nearly flat.

The present baryon mass spectrum, $P(k)$, in linear perturbation approximation, is the product of the transfer function and the assumed initial spectrum $\propto k^n$. The mass autocorrelation function, $\xi(r)$, is the Fourier transform of P , and the integral of $\xi(r)$ is

$$J_3(r) = \int_0^r \xi(r') r'^2 dr' \\ = 4\pi \int_0^\infty \frac{dk}{k} P(k) [\sin kr - kr \cos kr]$$

The velocity autocorrelation function^{1,2,12} is

$$v_c(r)^2 = 4\pi (Hf(\Omega))^2 \int_0^\infty dk P(k) \sin(kr)/(kr) \text{ where } f \approx \Omega^{0.6}$$

If the large-scale galaxy distribution traces mass then we can normalize $P(k)$ from J_3 for galaxies¹³:

$$J_3(50 \text{ Mpc}) = 6,200 \text{ Mpc}^3, H = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$$

Given n , Ω , and the model for ionization history this fixes $v_c(r)$. The results in Fig. 2 assume the matter is fully ionized. The label is the density parameter, Ω (with $h = 0.5$). The dashed curves correspond to $n = 0$, the solid curves to $n = -1$. A rough lower bound on n from the isotropy of the microwave background on scales of a few degrees is $n \geq -1$ (ref. 14). The tail of $v_c(r)$ extending to large r is dominated by the peak in $P(k)$, as is verified by the fact that if the peak is clipped $v_c(r)$ goes to zero at $r \sim 100$ Mpc.

Figure 3 shows how the results depend on the ionization history (for fixed $\Omega = 0.4$, $n = 0$). The parameter is $x_{\min}$, where the fractional ionization is the larger of thermal equilibrium and $x_{\min}$. In this and the other choices of Ω and n , $v_c(r)$ is not greatly different at $x_{\min} = 1$ and 0.1 , but at $x_{\min} = 0.01$ the peak in the transfer function is almost gone and the tail of $v_c(r)$ correspondingly reduced.

Now compare Fig. 2 to the required velocity field^{1,2,12}. The lowest curve, for $\Omega = 0.1$ and $n = 0$, is below the standard estimates. The case $\Omega = 0.1$, $n = -1$ gives a reasonable $v_c(0)$ but the coherence length may even be too large. The coherence length and r.m.s. velocity are both in reasonable accord with the usual

interpretation of the redshift data if $\Omega = 0.4$, $0 \geq n \geq -1$ and $x_{\min} \geq 0.1$. At $\Omega = 1$ the coherence length may be uncomfortably small.

The model can produce a velocity field that seems reasonable, and this success encourages further study of the model. The large velocity coherence length is the result of the prominent peak in the power spectrum of the mass distribution, thus suggesting that the presence of such a peak in the observations of the large-scale galaxy distribution should be investigated. The model predicts galaxy formation at very high redshift. One might hope for an observational test to distinguish between this and the very late galaxy formation in the standard adiabatic picture⁸ even in advance of the Hubble Space Telescope. On the theoretical side, it remains to be seen whether it will be possible to rationalize the two non-standard elements of the model, a density parameter that may be less than unity and a strongly inhomogeneous initial baryon number distribution.

This research was supported in part by the NSF.

Received 29 January; accepted 26 February 1987.

1. Vittorio, N., Juszkiewicz, R. & Davis, M. *Nature* **323**, 132–133 (1986).
2. Bond, J. R. *Galaxy Distances and Deviations from Universal Expansion* (eds Madore, B. F. & Tully, R. B.) 255–263 (Reidel, Dordrecht, 1986).
3. Wilkinson, D. T. *Science* **232**, 1517–1522 (1986).
4. Sandage, A. R. *Astrophys. J.* **307**, 1–19 (1986).
5. Rubin, V. C., Ford, W. K., Thonnard, N., Roberts, M. S. & Graham, J. A. *Astr. J.* **81**, 687–737 (1976).
6. Collins, A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506–508 (1986).
7. Burstein, D. et al. *Galaxy Distances and Deviations from Universal Expansion* (eds Madore, B. F. & Tully, R. B.) 123–130 (Reidel, Dordrecht, 1986).
8. Frenk, C. S., White, S. D. M., Efstathiou, G. & Davis, M. *Nature* **317**, 595–597 (1985).
9. Peebles, P. J. E. *The Early Universe* (ed. Unruh, W. G.) (Reidel, Dordrecht, 1986).
10. Peebles, P. J. E. *The Origin and Evolution of Galaxies* (eds Jones, B. J. T. & Jones, J. E.) 143–169 (Reidel, Dordrecht, 1983).
11. Peebles, P. J. E. *Astrophys. J.* **248**, 885–897 (1981).
12. Clutton-Brock, M. & Peebles, P. J. E. *Astr. J.* **86**, 1115–1119 (1981).
13. Davis, M. & Peebles, P. J. E. *Astrophys. J.* **267**, 465–482 (1983).
14. Peebles, P. J. E. *Astrophys. J.* (in the press).

The infrared and radio-continuum emission of the galactic disk

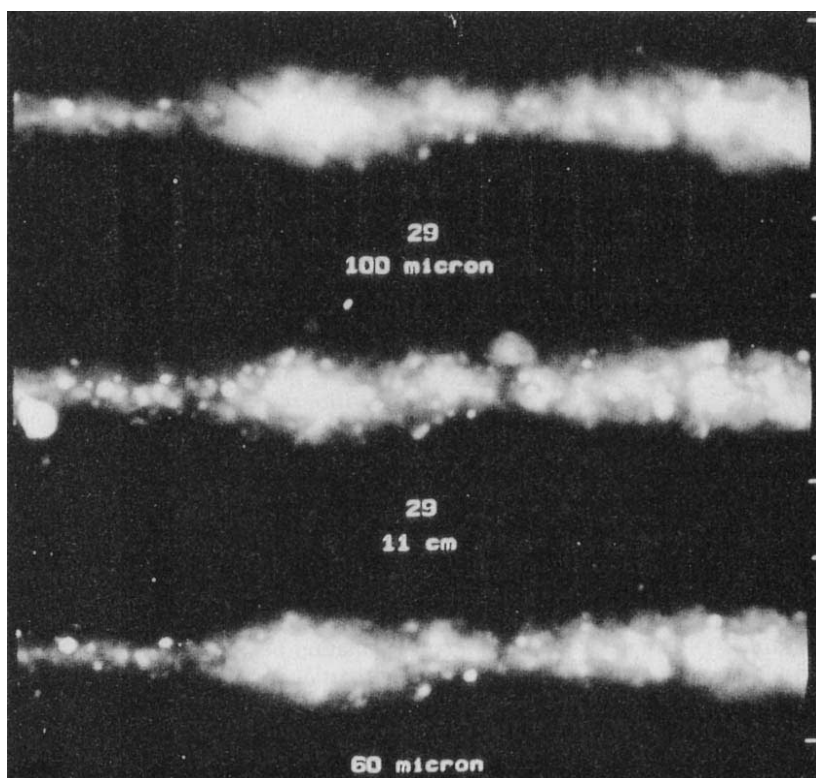
C. G. T. Haslam* & J. L. Osborne†

* Max-Planck-Institute für Radioastronomie, Auf dem Hügel 69, D-5300 Bonn 1, FRG

† Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

In 1983 the Infrared Astronomical Satellite (IRAS) surveyed the sky in four wavebands in the far infrared (FIR). One of the products of this survey¹ is a detailed picture with an angular resolution of ~ 4 arc min of the infrared emission of the interstellar dust in the galactic plane. Here we point out the remarkably detailed correlation between the 60- μm band emission and most of the features in the 11-cm radio-continuum emission as measured by Reich *et al.*² with a beam width of 4.3 arc min. This correlation has two consequences. First, it gives support to the model of Cox *et al.*³ in which the 60- μm band emission arises mainly from dust associated with extended low-density (ELD) H II regions. Second, it provides a way of separating the thermal component of the radio continuum, identified by its associated infrared emission, from the synchrotron component. When the infrared and radio maps are compared, the known supernova remnants stand out boldly as radio sources which do not have the corresponding infrared emission. Candidates for hitherto undiscovered supernova remnants can be found by looking for additional radio sources lacking infrared emission. Larger-scale structure of the galactic magnetic field near the galactic plane will be revealed after the synchrotron emission, uncontaminated by thermal emission, has been obtained in this way.

Fig. 1 Grey-scale representations of the distributions of intensity of far infrared and radio continuum emission in the galactic plane. Each of the three strips covers the range of galactic longitude from 23° (right) through 29° (centre) to 35° (left) and galactic latitude from -1.5° to $+1.5^\circ$. The top and bottom strips show IRAS infrared intensities in the $100\text{ }\mu\text{m}$ and $60\text{ }\mu\text{m}$ bands, respectively. Centre strip shows the 11-cm radio-continuum intensities of Reich *et al.*².

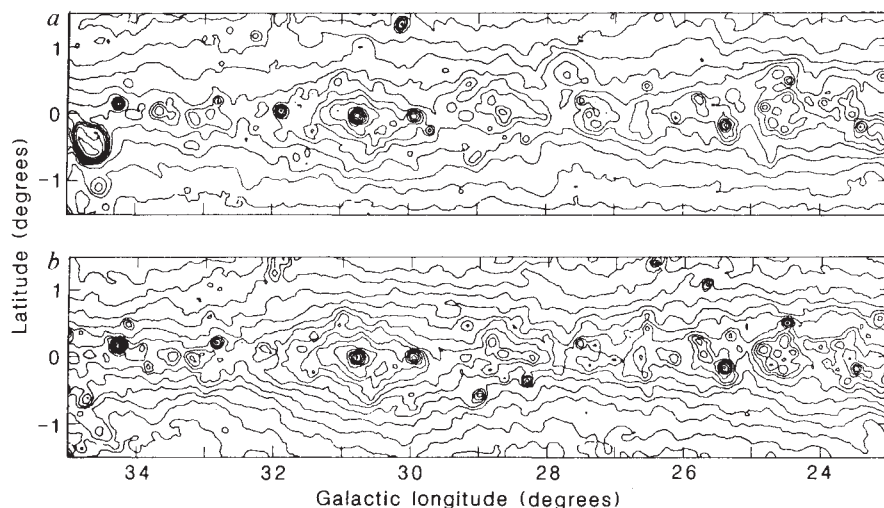


There is general agreement that the diffuse FIR emission from the galactic disk comes from interstellar dust which is heated by the absorption of starlight. The location of the dust grains and the sources of the interstellar radiation that contribute most to the heating are a matter of debate. In one model⁴, most recently modified by Cox *et al.*³ much of the diffuse emission comes from 'warm' dust at 30 to 40 K. This dust is mainly associated with the ionized gas in ELD H II regions which has been ionized by observable O stars which spend most of their life outside the clouds of their birth. Typical densities of the ELD regions are 5 to 10 electrons per cm^3 . The diameters of the regions range from 20 to 250 pc. A smaller fraction of this warm dust is associated with the molecular gas surrounding B stars. A second model⁵ has the diffuse emission coming from dust particles permeating the giant molecular clouds (GMC). The dust is assumed to be heated by embedded, and thus unobservable, O and B stars. By contrast, in the ELD model the dust in the molecular clouds is heated to only $\sim 14\text{ K}$ by the general interstellar radiation field attenuated by absorption and the contribution of this dust to the $60\text{ }\mu\text{m}$ infrared emission is negligible. In the ELD model the infrared emission would be expected to correlate well with the thermal continuum emission from the H II regions. In the GMC model the correlation would be with the ^{12}CO column density which acts as a tracer of the molecular gas.

Maihara *et al.*⁶ concluded from the results of a balloon flight survey at $150\text{ }\mu\text{m}$ with an angular resolution of about 1 degree that the overall similarity between their contour map and the 5 GHz contour map of Altenhoff *et al.*⁷ supported the ELD model. By contrast Hauser *et al.*⁸ maintained, from their balloon flight survey between 150 and $300\text{ }\mu\text{m}$ with angular resolution 10 arc min, that there was a better correlation with the ^{12}CO intensity profiles of Cohen *et al.*⁹ than with a 5-GHz radio-continuum profile of the same angular resolution. The reason that these different conclusions can be arrived at is that the giant H II regions are generally found to be associated with these clouds. This is to be expected as H II regions result from star formation which is enhanced in regions occupied by giant

molecular clouds. We would claim, however, that with the improved angular resolution of the present data it has become clear that the primary correlation, at least in the $60\text{-}\mu\text{m}$ band, is between the infrared and the radio continuum. To illustrate this we show in Figs 1 and 2 that part of the sky is within 1.5° degrees of the galactic plane between galactic longitudes of 23° and 35° . In Fig. 1 the top strip shows the $100\text{-}\mu\text{m}$ band intensities, the middle the 11-cm radio intensities, and the bottom the $60\text{-}\mu\text{m}$ band intensities. The linear grey scales have been matched approximately by eye. Figure 2 allows a more quantitative comparison to be made of the intensities. Here the $60\text{-}\mu\text{m}$ intensities are those remaining after the removal of the, relatively small, contributions of zodiacal light and emission from dust associated with neutral atomic hydrogen (H I). The brightness temperatures from the 11-cm published survey² were measured with respect to an arbitrary base level and we have estimated its value. The method of obtaining these corrections is given below. One sees that not only do the discrete H II regions, which are bright infrared sources, appear prominently in the radio, but the much more extended infrared complexes also show radio emission that mirrors their form in considerable detail. Although the correlation between the radio and the $100\text{-}\mu\text{m}$ intensities is good, the correlation with the $60\text{-}\mu\text{m}$ intensities appears slightly better. We would attribute this to the effect of the dust associated with the H I gas. In the model of Cox *et al.*³, of the mean intensity at $60\text{ }\mu\text{m}$ within 1° of the galactic plane, 82% is from dust associated with H II and the remainder is from dust associated with H I. At $100\text{ }\mu\text{m}$ the H I dust, being cooler than the H II dust, accounts for about 40% of the emission from the same region. At $200\text{ }\mu\text{m}$, appropriate to the survey of Hauser *et al.*⁸, about 14% of the total flux is predicted to come from dust associated with molecular clouds while the remainder is contributed to in equal parts by the H I and H II associated dust. When one compares the infrared and radio fluxes of the discrete H II regions (that is, those with an angular size $< 4\text{ arc min}$) there is a scatter in the ratios but the values are within a factor of 2 either way of the value that is found for the diffuse features. Although the gas densities and other conditions may differ

Fig. 2 Contour maps of *a*, the 11-cm intensities and *b*, the 60- μ m intensities for the same region as for Fig. 1. The contribution of zodiacal light and H I associated dust has been subtracted from the 60 μ m intensities. A base-level correction of 660 mK has been added to the 11 cm brightness temperatures. The contour levels increase by a factor of $\sqrt{2}$ from one to the next. The outer contours on the 60- μ m and 11-cm maps are at 50 MJy sr $^{-1}$ and 710 mK brightness temperature respectively. At a wavelength of 11 cm the conversion factor from brightness temperature to intensity is 0.22 MJy sr $^{-1}$ K $^{-1}$.



considerably between the diffuse and compact H II regions it is seen that the relative production of infrared and radio emission remains reasonably constant.

One obvious difference between the radio and infrared pictures in Fig. 1 is the bright blob of radio emission at galactic longitude, $l = 34.7^\circ$. This is the bright supernova remnant W44. Examination of the infrared maps shows no distinguishable emission from this source. The dust within the remnant appears to be negligible in comparison with that in an H II region of comparable brightness at 11 cm. All of the catalogued supernova remnants in this region show the same signature of radio emission without significant associated infrared emission. This is not at variance with the measured infrared emission from some supernova remnants, favourably situated for its detection in the outer parts of the Galaxy or at relatively high galactic latitude. The Cygnus Loop¹⁰ has a 60- μ m flux of 1,430 Jy for example, giving a ratio of 60- μ m to 11-cm flux of 11. The Crab nebula¹¹ has a ratio of 0.3. This is to be compared with ratios ≥ 500 for individual H II regions and 1,300 for our overall fit to the diffuse regions (see below). For the catalogued supernova remnants in the region we have studied all the ratios, most of which are upper limits, are < 20 . This suggests a simple method for finding hitherto undiscovered supernova remnants near the galactic plane. The current ways of distinguishing H II regions from supernova remnants rely on radio spectral index determination and measurements of polarization and recombination line emission. As well as involving additional detailed observations each of these tests can give uncertain results. The spectral index, α , where the flux at frequency f is proportional to $f^{-\alpha}$ has the value of 0.1 for H II regions. Although a 'typical' remnant has $\alpha = 0.7$, some have indices as low as 0.1. A search based on a spectral index may miss flat spectrum remnants. H II regions should be distinguishable from supernova remnants in having recombination line emission but no radio polarization. Both of these tests can lead to ambiguous results for sources near to the galactic plane because of emission from other sources along the line of sight. Reich *et al.*¹² have concluded, for example, as the result of such tests, that the radio source that can be seen at $l = 30.7^\circ$, $b = +1.0^\circ$ is a supernova remnant. The fact that there is no corresponding source on the infrared pictures confirms this. On the basis of the spectral index of its integrated flux the large radio object at $l = 27.7^\circ$, $b = +0.6^\circ$ has been classified as a filled centre supernova remnant¹³. Its absence from the infrared pictures confirms that the whole of it is non-thermal. By the same criterion the radio source at $l = 31.7^\circ$, $b = -0.7^\circ$ is also a supernova remnant candidate.

A further quantitative comparison between the 60- μ m and 11-cm intensities is given in Fig. 2 where the profiles along the

galactic plane over a range of 40° in longitude of the intensities averaged over $\pm 0.5^\circ$ in latitude are shown. Peaks which are all or in part due to catalogued supernova remnants are marked. Apart from these there is a strong correlation between the two profiles. If comparison is made with the corresponding ^{12}CO profile (see, for example⁸) one sees that the maxima around $l = 24^\circ$ and 31° are present but there is no such detailed correlation.

The summed emission from recognizable supernova remnants gives only a small contribution to the total synchrotron emission, the bulk of which is produced by cosmic-ray electrons interacting with the large-scale galactic magnetic field. In the past, the separation of thermal and synchrotron radio emission on a large scale has been made by comparing radio surveys at low frequencies (typically 408 MHz) with those at 2.7 GHz (11 cm) or 5 GHz. This requires the spectral indices of the two components to be known. An index $\alpha = 0.1$ is to be expected for the thermal emission but the index of the synchrotron emission may vary from place to place from its usually adopted value of 0.7. There is an additional difficulty in determining the base level of the high frequency survey with sufficient accuracy to be able to rely on the observed spectral index. From the investigations that have been made one knows that the summed radio synchrotron emission from the galactic disk at 11 cm is comparable with the thermal emission. The scale height of the former is several times greater than that of the latter, however, so that thermal emission dominates close to the plane. If one has more reliable and detailed separation of the components the modulation of the synchrotron emission by the spiral structure of the Galaxy can be investigated with greater confidence. For instance, steps in the profile of the radio-continuum emission along the galactic equator have long been associated with spiral structure but because of the contribution from regions of thermal emission which also lie along the spiral arms it has been difficult to determine the true degree of enhancement of synchrotron emission in the spiral arms. We propose the following procedure to separate the thermal emission from the synchrotron emission. First, remove the zodiacal light contamination from the 60- μ m data. At high galactic latitudes the zodiacal light dominates in this waveband but within 1.5° of the plane it represents only a small correction. Second, remove that component of the 60 μ m emission which comes from dust associated with the H I gas. At galactic latitudes greater than 10° the H I associated dust is negligible and an empirical relationship between 60- μ m intensity and column density of H I has already been found. We shall assume that the same relationship holds at lower latitudes. In any case this is a minor contribution. Third, find an empirical relation between the remaining 60 μ m emission and the thermal

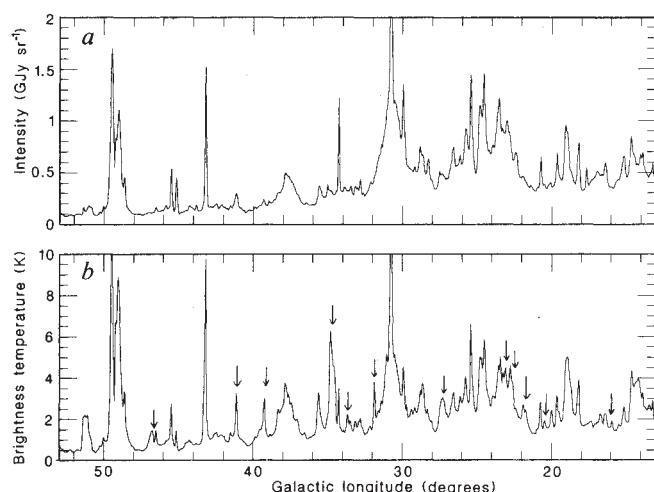


Fig. 3 Profiles of *a*, 60- μ m band infrared emission and *b*, 11 cm radio-continuum emission over a range of 40° of galactic longitude. Each profile gives the intensity averaged over $\pm 0.5^\circ$ of galactic latitude. The small arrows on the radio profile indicate the positions of catalogued supernova remnants.

part of the 11-cm emission. This is done by plotting the 11-cm intensity against the net 60- μ m intensity pixel by pixel. It is found that there is a well-defined lower envelope to this scatter plot to which a straight line can be fitted. We interpret this as showing that the infrared emission always has thermal radio emission associated with it (except for individual stars which can readily be identified by their showing maximum brightness in the 12- μ m band) and that points on the lower envelope are for directions where the non-thermal emission is negligible. A preliminary linear fit to this lower envelope indicates an intensity ratio of 1,300:1 between the 60- μ m H II associated intensity and the thermal component of the 11-cm intensity. The intercept on the 11-cm intensity axis indicates that a base-level correction of 660 mK should be added to the 11-cm brightness temperatures. The brightness temperatures given in ref. 2 are measured with respect to an arbitrary base level and this correction lies within the expected range. Finally, use this empirical relation to subtract the thermal component from the 11-cm maps pixel by pixel.

This procedure is now in progress. The 11-cm survey covers the longitude range from 358° to 76°. The survey of Haynes *et al.*¹⁴ at 5 GHz, which has a similar angular resolution, is available for longitudes 190° to 40°. Together they will allow coverage of practically the whole of the inner half of the Galaxy.

We thank the IPMAF team of Rutherford Appleton Laboratory for their help in providing the IRAS data and the software with which to analyse them. We thank Miss A. Broadbent for providing some of the unpublished results of her work on this topic.

Is there dust in galactic haloes?

J. Mayo Greenberg*, Federico Ferrini*†, Bruno Barsella‡ & Santi Aiello‡

* Laboratory Astrophysics, Rijksuniversiteit Leiden, 2333 CA, The Netherlands

† Instituto di Astronomia, Università di Pisa, Pisa 56100, Italy

‡ Dipartimento di Fisica, Università di Firenze, Firenze 59125, Italy

The ubiquitous presence of dust within the disks of spiral galaxies is well established. Dust is present in large quantities over enormous extents in the Universe; for example, Burton¹ concludes that, in our Galaxy, dust is distributed very similarly to the H I distribution, with an exponential radius in the plane of 8 kpc and an exponential scale height perpendicular to the plane of 120 pc. Simple dynamical arguments about the motion of interstellar grains out of the galactic disk show the possibility of confining a significant quantity of dust in a potential well high above the galactic plane. Although the possibility of dust being repelled away from the galaxy by the radiation pressure of the stars in the disk has earlier been considered², we have noted that the functional difference between the gravitational and radiation forces, as the dust recedes, may lead to the existence of stable regions outside the galaxy reminiscent of the libration points for interplanetary dust produced by competing gravitational forces alone (refs 3, 4; F.F., J.M.G., B.B. and S.A., in preparation). Here we predict that the presence of dust in these regions may be revealed in bright edge-on galaxies, especially by using the polarization of the scattered light from the symmetric lanes. The detection of scattered light above the galactic plane may be an indicator that the parent galaxy has not suffered close encounters with other galaxies at least within the timescale required to establish the dust layers.

We consider the forces acting on the dust which is located near the upper and lower faces of the disk of a spiral galaxy. In this position the radiation force coming from the stars of the galaxy is not isotropic and results in a net effect of expelling dust from the disk.

Here we present only the physical scheme of the calculation along with selected results of some detailed calculations. A fuller description will be given elsewhere (F.F., J.M.G., B.B. and S.A., in preparation). To determine the radiation force on the dust grains, we need to know the light distribution of the model galaxy due both to the disk and bulge components. Given the optical properties of the grains, we derived the effective cross-section for radiation pressure, Q_{pr} . Light from galaxies is generally given as a surface brightness distribution, integrated over frequency, so to obtain the effective cross-section integrated over the spectrum, we convolved the Q_{pr} with several black-body spectra $B_\nu(T)$ (assuming that the radiation of a galaxy could be roughly approximated in this way) at various effective temperatures (ranging from 3,000 K to 12,000 K). We define:

$$\bar{Q}_{pr} = \int Q_{pr\nu} B_\nu(T) d\nu / \int B_\nu(T) d\nu \quad (1)$$

We considered the effect of the radiation coming from the disk of the galaxy on the grains, whose geometrical (spheres ranging from 0.01 μ m to 0.25 μ m radius) and optical (schematized as pure silicate or pure graphite) properties are varied. We determined, in other words, the mean radiation force on the grains, ignoring for the moment the stochastic term which can result from random single energetic photon collisions. A fully dynamical treatment should include this effect.

The other fundamental force acting on the grain is gravity. To calculate its effect we must know the mass distribution of the galaxy in its various components: disk, bulge and halo. We chose NGC3198 as a model galaxy, one for which both extended

Received 8 January; accepted 18 March 1987.

1. IRAS Explanatory Supplement (eds Beichman, C. A., Neugebauer, G., Habing, H. J., Clegg, P. E. & Chester, T. J.) (JPL D-1855 Jet Propulsion Laboratory, Pasadena, 1985).
2. Reich, W., Fürst, E., Steffen, P., Reif, K. & Haslam, C. G. T. *Astr. Astrophys. Suppl. Ser.* **58**, 197-248 (1986).
3. Cox, P., Krügel, E. & Mezger, P. G. *Astr. Astrophys.* **155**, 380-396 (1986).
4. Mezger, P. G., Mathis, J. S. & Panagia, N. *Astr. Astrophys.* **105**, 372-388 (1982).
5. Fazio, G. G. & Stecker, F. W. *Astrophys. J.* **207**, L49-L52 (1976).
6. Maihara, T., Oda, N. & Okuda, H. *Astrophys. J.* **227**, L129-L132 (1979).
7. Altenhoff, W. T., Downes, D., Goad, L., Maxwell, A. & Rinehart, R. *Astr. Astrophys. Suppl. Ser.* **1**, 319-355 (1970).
8. Hauser, M. G. *et al.* *Astrophys. J.* **285**, 74-88 (1984).
9. Cohen, R. S., Cong, H., Dame, T. M. & Thaddeus, P. *Astrophys. J.* **239**, L53-L56 (1980).
10. Braun, R. & Strom, R. G. *Astr. Astrophys.* **164**, 208-217 (1986).
11. Marsden, P. L. *et al.* *Astrophys. J.* **278**, L29-32 (1984).
12. Reich, W., Fürst, E., Reich, P., Sofue, Y. & Handa, T. *Astr. Astrophys.* **155**, 185-192 (1986).
13. Reich, W., Fürst, E. & Sofue, Y. *Astr. Astrophys.* **133**, L4-L7 (1984).
14. Haynes, R. F., Caswell, J. L. & Simons, L. W. J. *Aust. J. Phys. Astrophys. Suppl.* **45**, 1-87 (1978).

optical and radio observations at large radial distances were available as well as its mass distribution. The rotation curve of this galaxy is well known up to large distances. For the light distribution we followed Wevers⁵, and for the mass distribution we used the maximum disk model of van Albada *et al.*⁶.

We summed the resulting forces at various positions perpendicular to the galactic plane (axial symmetry), for several different compositions and radii of the grains. The resulting vectorial distribution of the force is obviously not central, and shows an interesting behaviour (see Fig. 1 for some typical situations). Along a radial line starting from the centre of the model galaxy, the net force at first expels the grains but, further out, becomes attractive, resulting in a potential well which may confine the grains around a certain distance.

A simple order of magnitude argument can explain the reason for this result. Let us compare the ratio of radiation force to gravitation on a grain of radius a , mass m , radiation pressure effective cross section $\bar{Q}_{pr}$ which is located at a certain distance, r , from the galactic centre and 45° above the plane. The galaxy (NGC3198) is a disk with optical radius $R_L = 18$ kpc, mass radius $R_M = 50$ kpc, surface luminosity distribution, $I_0 e^{-\gamma\rho}$, surface mass distribution $\sigma_0 e^{-\gamma\rho}$, halo mass $M(r)$, where $I_0 = 190 L_\odot/\text{pc}^2$, $\sigma_0 = 687 M_\odot/\text{pc}^2$, $\gamma = 1/2.68$ kpc and $M(r) = 10^{10} M_\odot[(r/2) - 3]$ for $r \geq 8$ kpc (r in kpc).

The radiation force is:

$$F_{\text{rad}}\left(r, \frac{\pi}{4}\right) = \frac{\pi a^2 \bar{Q}_{pr}}{c} \int_{S_{\text{disk}}} I_0 e^{-\gamma\rho} \frac{\mathbf{r} - \mathbf{\rho}}{|\mathbf{r} - \mathbf{\rho}|^3} dS \quad (2)$$

A reasonable estimate of the integral is made by noting that most of the contribution comes from the central area of radius $\gamma R_L \approx 6$ kpc. Thus

$$F_{\text{rad}} \approx 1.4 \frac{a^2 \bar{Q}_{pr}}{c} I_0 \frac{R_L^2}{r^2} \quad (3)$$

For the gravitational force we see from van Albada *et al.*⁶ that most of the disk mass is internal to 9 kpc, which we assume to be an effective radius, R_M , of the mass distribution:

$$F_{\text{grav,disk}} \approx -0.8 G m \sigma_0 \frac{R_M^2}{r^2} \quad (4)$$

$$F_{\text{grav,halo}} = -G m \frac{M(r)}{r^2} \quad (5)$$

Equating the net force to zero for various grains parameters as, for example:

silicate	$a = 0.05 \mu\text{m}$	$\bar{Q}_{pr} = 0.28$
silicate	$a = 0.15 \mu\text{m}$	$\bar{Q}_{pr} = 0.55$
graphite	$a = 0.02 \mu\text{m}$	$\bar{Q}_{pr} = 0.095$

may be readily seen to give an 'equilibrium' radius in the range 15–20 kpc.

This simple demonstration of the force field shows the possibility of confining dust material high above the galactic plane. Naturally, to have a complete idea of what could happen to the dust at high galactic latitude, we must integrate the equations of motion for the grain, that is solve for the orbits taking into account: stochastic single photon collisions, possible drag on the grains by the gas material present in the halo, and the Poynting–Robertson effect.

Another possibility which remains to be considered in detail is that, if the particles are charged, the effects of magnetic forces may become significant. It is difficult to assess this accurately because of lack of precise information on magnetic field strengths outside the galactic plane. The mean magnetic field within spiral arms in our galaxy is $\sim 2\text{--}3 \times 10^{-10}$ tesla. Above the galactic plane (assuming a coupling of the field to the gas) the magnetic field may decrease exponentially with a scale height of several hundred parsecs⁷. At a distance z above the plane of 250 pc the Lorentz force exerted by a 10^{-10} tesla field on a singly

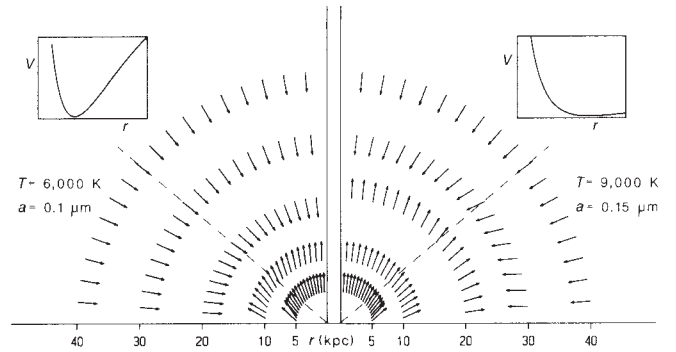


Fig. 1 Vectorial distribution of combined gravitational and continuous radiation forces on silicate particles above the galactic plane. Results shown are for particle radii 0.1 and 0.15 μm and galactic (black-body) radiation temperatures of 6,000 K and 9,000 K respectively. Upper diagrams are the potential fields at elevation angle 45° .

charged 100 nm particle moving at 1 km s^{-1} is less than 1/5 that of the radiation pressure out to radial distances of about 5 kpc. Even if the field remains at 10^{-10} tesla out to a distance of 5 kpc, the Lorentz force may be only comparable with the radiation force. Thus, although a regular magnetic field parallel to the galactic plane may confine 1 km s^{-1} grains, random magnetic fields would tend to reduce this effect. Finally, the decoupling of a grain from the magnetic field each time it becomes neutral allows the radiation repulsion to take over again and, although this occurs sporadically, it is monotonic in direction.

Before presenting detailed computer calculations which confirm the dynamical confinement effect by an effective minimum potential at some kiloparsecs from the galactic plane, we wish to address first the very basic question: is it possible to observe this dust if it is indeed suspended above the galactic plane? The dust would certainly be invisible in the infrared, because the grain temperature is too low, certainly less than the 10–15 K temperatures of grains within the plane and probably approaching the universal black-body temperature of 3 K⁸. However, it may be possible to observe the scattering effect of this material in the visible. Here we present a schematic calculation of the ratio of scattered light density to the surface brightness density of the galaxy. The infrared emission by extremely small—5 nm particles such as the polycyclic aromatic hydrocarbons⁹ has not yet been considered because although they may be expelled from the galaxy it is not obvious how to calculate the radiation pressure on such non-classical scatterers. We are investigating this further.

The amount of light per unit surface scattered by a grain distribution with a surface density N is

$$\frac{N \pi a^2 Q_{sc}}{4\pi} \int_{S_{\text{disk}}} I_0 e^{-\gamma\rho} \frac{\mathbf{r} - \mathbf{\rho}}{|\mathbf{r} - \mathbf{\rho}|^3} dS \quad (6)$$

We estimate the integral as we have done above for the radiation force and divide by the maximum surface brightness of the disk I_0 , considering the dust layer at a distance of the order of R_L . The surface brightness ratio (SBR) becomes:

$$\text{SBR} \approx \frac{N a^2 Q_{sc}}{20} \quad (7)$$

To estimate N we fix a simple geometry for the dust particle distribution. It is easy to show (see F.F. *et al.*, in preparation) that the results do not depend strongly on this detail. Let us suppose the configuration acquired by the dust is a torus, with radius R and cross-section t . Then, the surface particle density is related to the volume particle density n by $N = \pi/2 (tn)$ where n is proportional to the total dust mass in the galaxy (M_{dust}), divided by the grain mass and the torus volume, multiplied by

the fraction of dust expelled from the galactic disk (f_x). Thus equation (7) becomes

$$\text{SBR} \approx 10^{-3} f_x Q_{\text{sc}} \frac{M_{\text{dust}}}{R t a} \quad (8)$$

To evaluate this ratio, we use a rather conservative estimate for the geometrical parameters; that is, we put $R = R_L$ and $t = R_L/3$; $M_{\text{dust}} \approx 10^{-3} M_{\text{gal}}$. So that with a in μm we get:

$$\text{SBR} \approx 0.3 \frac{Q_{\text{sc}} f_x}{a} \quad (9)$$

We estimate f_x by considering that a small layer (about 10 pc thin) at two scale heights above the galactic plane has a high probability of undergoing radiation pressure repulsion. Then $f_x \geq 10^{-2}$ and

$$\text{SBR} \approx 3 \times 10^{-3} \frac{Q_{\text{sc}}}{a} \quad (10)$$

For a wavelength of 4,000 Å, and a particle radius of 0.01 μm , $Q_{\text{sc}} \approx 10^{-4}$ which gives $\text{SBR} \approx 3 \times 10^{-5}$. But, if the grains have a radius of 0.1 μm , then $Q_{\text{sc}} \approx 1$ and $\text{SBR} \approx 3 \times 10^{-2}$. In the former case the ratio of scattered light to the galactic light is down by ~ 11 mag, while in the second case it is down by only 4 mag.

The polarization by the 0.1 μm particles at 90° is about 20% if they are anything like those in the Milky Way¹⁰. This amount of polarization should make the detection of the displaced

galactic dust much easier, even if the ratio of its surface brightness to the galaxy is low. In the most advantageous case of an edge-on galaxy the dust lanes would show up as symmetrical lanes, that is, geometrically correlated and similarly aligned, and the regions of brightness and polarization would both be parallel to the plane of the galaxy.

The expulsion of dust from galactic disks may have a significant effect on the chemical evolution of galaxies¹¹ if, in the past, relatively more of the available Si, Mg and Fe were found in the dust than C, N, O as is the case at the present time.

F.F. gratefully acknowledges a grant from the ZWO (Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek) and CNR which supported a visit to Huygens Laboratorium in Leiden, where this work was completed.

Received 24 October 1986; accepted 18 March 1987.

1. Burton, W. B. in *Proc. C.I.S.I.F. Int. School of Physics "Enrico Fermi"* Varenna (eds Greenberg, J. M. & Aiello, S.) (Academic, New York, 1986).
2. Chiao, R. Y. & Wickramasinghe, N. C. *Mon. Not. R. astr. Soc.* **159**, 361–373 (1972).
3. Burns, J. A., Lamy, P. & Soter, S. *Icarus* **40**, 1–48 (1979).
4. Mukai, T. *Astr. Astrophys.* **153**, 213–217 (1985).
5. Wevers, B. M. H. R. thesis, Groningen Univ. (1984).
6. van Albada, T. S., Bahcall, J. N., Begeman, K. & Sancisi, R. *Astrophys. J.* **295**, 305–313 (1985).
7. Harnett, J. I. & Reynolds, J. E. *Mon. Not. R. astr. Soc.* **215**, 247–257 (1985).
8. Greenberg, J. M. *Astr. Astrophys.* **12**, 240–249 (1971).
9. Léger, A., d'Hendecourt, L. & Boccaro, N. (eds) *Polycyclic Aromatic Hydrocarbons and Astrophysics* (Reidel, Dordrecht, 1986).
10. Greenberg, J. M. & Hanner, M. S. *Astrophys. J.* **161**, 947–959 (1970).
11. Matteucci, F. *Astrophys. J.* **305**, L81–L84 (1986).

Long-term changes in the tropical Pacific surface wind field

Kim D. B. Whysall*, Neill S. Cooper & Grant R. Bigg*

The Hooke Institute for Atmospheric Research,
Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, UK

There has been considerable research to determine climate variability on the decadal timescale, primarily using temperature data^{1–11}. In the tropical Pacific the dominant climate fluctuation is the El Niño–Southern Oscillation phenomenon^{12–14}, which includes large anomalies in the wind field. The surface wind velocity is the meteorological variable most widely, and reliably, observed at sea and hence we have performed an analysis of sixty-four years of Pacific wind data. Here we report that the results show that during the period 1940–44 there was a strong westerly anomaly, of the order of 1 m s^{-1} in a climatology of 5 m s^{-1} . This anomaly was concurrent with anomalies in sea surface temperature¹ and Pacific sea-level pressure². During this period the warming of the ocean and the westerly wind anomaly resemble a prolonged El Niño. We also show from the observations that, in the period 1950–81, there was a trend towards increasing trade winds over much of the Pacific Ocean. A similar trend has also been observed in Atlantic Ocean data^{3,15,16}.

Difficulties in making precise temperature measurements at sea result in uncertainties in the interpretation of long atmospheric and sea surface temperature time series. For example, ship-based observations of atmospheric temperature are liable to inaccuracies of over 1°C due to ship heating¹⁷. Alterations in sea surface temperature (SST) observational techniques, from uninsulated buckets to insulated buckets and engine-intake temperature, have introduced further instrumental biases^{1,2,8}. When such effects are removed the resulting data indicate that marine

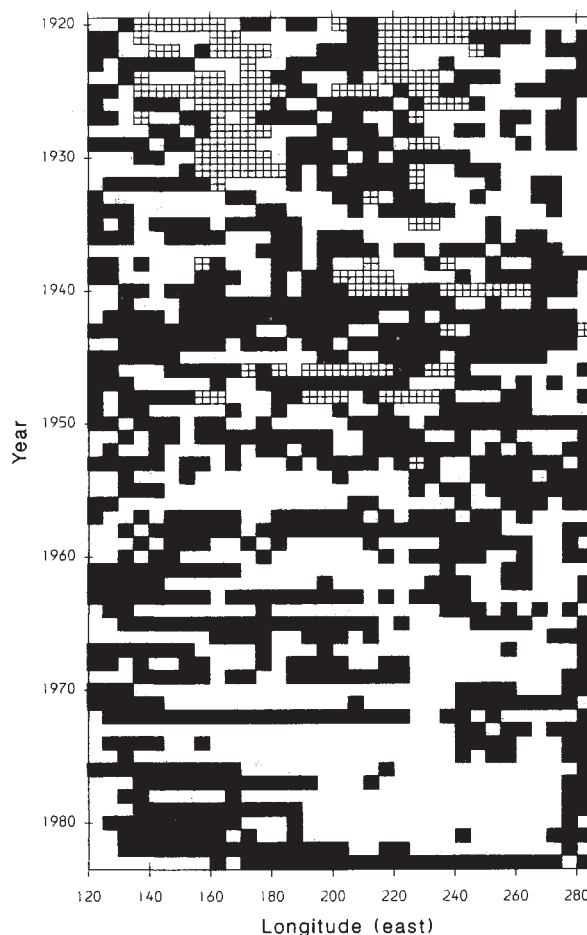
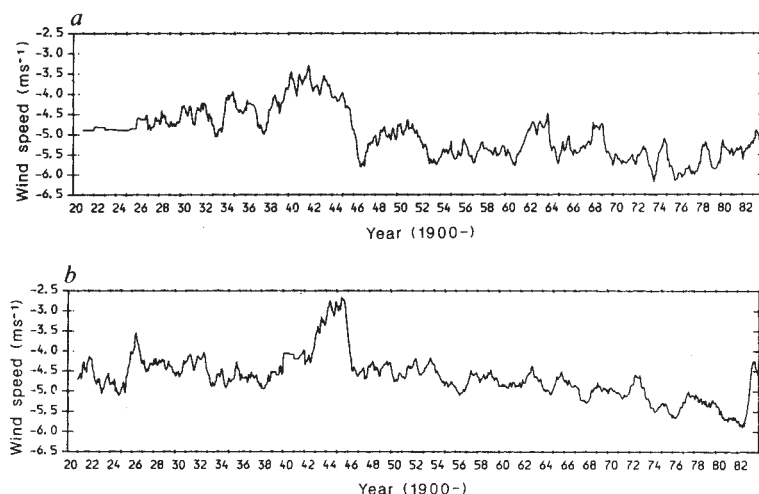


Fig. 1 Longitude-time plot of the zonal wind anomalies, annual average from 5°S to 5°N . Shaded regions represent positive (westerly) anomalies, unshaded regions negative anomalies while hatched regions contain no data. Note the basin-wide anomaly of 1942–1944 and the trend to more negative anomalies since 1950.

* Present addresses: Meteorological Office, London Road, Bracknell RG12 2SZ, UK (K.D.B.W.); School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK (G.R.B.).

Fig. 2 Time series of 12-month running mean zonal wind averaged over *a*, North Pacific (5°N – 25°N , 150°E – 130°W) and *b*, South Pacific (0°S to 20°S , 160°W to 90°W). Maximum standard errors are less than 0.1 m s^{-1} .



air temperature, global atmospheric temperature and, perhaps, SST, were anomalously high in the period 1935–1945^{1–11}. Also observed is a global atmospheric cooling of about 0.3°C between 1945 and 1970^{1,2,5–8}. As the atmospheric circulation is related to the temperature structure, concurrent wind changes will have occurred with these temperature variations, and the results presented here demonstrate this connection.

The data used in this study are ship winds from the comprehensively quality-controlled UK Meteorological Office Main Marine Databank. The data comprise the monthly mean wind, its standard deviation and the number of observations used for each of the 5° by 5° squares in the region 120°E to 75°W , 30°S to 30°N for the years 1920 to 1983. These values have not been smoothed in either space or time.

As with temperature observations, there is the risk not only of observational error but also of trends in such errors. In the case of wind observations this is due to the subjective nature of the observations, which are mainly visual, being based on the Beaufort Scale. There is, however, no indication of a spatially coherent trend in the wind data except in particular regions. The monthly standard deviation data show no trend and, while the density of shipping has increased, the shipping routes have remained essentially the same. Also there is no indication in the data of the change from visual to instrumental observation in the mid-1960s. Because of these properties we believe that the variations in the observed wind field presented in this paper were due to actual atmospheric changes rather than to observational biases.

The winds in the equatorial Pacific are known to be particularly sensitive indicators of the El Niño–Southern Oscillation (ENSO) phenomenon^{12–14}. During such events the eastern equatorial Pacific Ocean is anomalously warm by several degrees Celsius and the trans-Pacific atmospheric pressure gradient weakens leading to westerly anomalies in the surface wind field. This can be clearly seen in the longitude–time plot in Fig. 1 which shows the zonal velocity anomalies averaged for each year from 5°S to 5°N . The monthly anomaly is calculated with respect to the 64-year monthly climatology for each box, and then averaged over each calendar year. In Fig. 1 positive (westerly) anomalies are shaded, negative (easterly) anomalies are unshaded while hatching represents regions of no data. Thus, interesting climate anomalies can be observed in Fig. 1, although not their significance level. Westerly anomalies associated with ENSO events are evident in this diagram, in particular in 1957, 1958, 1963, 1965, 1972 and 1983. There are also other clear signals in the data, and it is these which are of interest in this paper. The most extreme event is the basin-wide westerly anomaly of 1942–44; these are the only years in the whole period analysed in which more than 80% of the ocean has a westerly

anomaly. (In comparison, during the ENSO event of 1982–1983 only 70% of the ocean had a westerly anomaly.) During the same period both the global SST and the global night-time marine air temperature were their warmest this century¹. The results presented here substantiate a hypothesis postulated by Barnett² that there was a multi-year climate event in the atmosphere–ocean system in the late 1930s through to the mid-1940s.

The amplitude of the anomaly can be seen in Fig. 2, which shows time series of 12-month running means of the wind in both the North and South Pacific trade-wind belts. The two regions used are chosen to correspond to the central areas of Barnett's¹⁸ first empirical orthogonal eigenfunction of the equatorial Pacific zonal wind field. To avoid sparse data having an unduly large influence, climatology, rather than the observed wind, was used in months with less than 10% of the average number of observations for that region.

From Fig. 2 it can be seen that the prolonged westerly anomaly was basin wide, with effects on the trade winds over a slightly longer period than that in the equatorial region. The amplitude of the effect is over 1 m s^{-1} , and there was a dramatic return to stronger trades in 1946. The spatial structure of the anomaly can be seen in Fig. 3 which shows the 64-year wind climatology for 1920–83 obtained by using all the data and also the difference from the mean field for the period of 1940–44. The wind is observed to have decreased by about 1 m s^{-1} over most of the tropical Pacific. The coherent structure of the anomaly indicates that it is not a chance fluctuation but a true climate anomaly. This suggests a major collapse of the tropical Pacific trades which is consistent with the observed increases in the SST and air temperature^{1–11}, since as the ocean surface warms so does the atmospheric boundary layer and hence, by the Clausius–Clapeyron relationship between saturated vapour pressure and temperature¹⁹, evaporation decreases. This results in a positive feedback mechanism. It is not clear which of these phenomena occurred first nor why the westerly anomaly was so prolonged. Further research is required to determine the cause and persistence of this major climate anomaly. The grouping of El Niños in the 1880s, indicated by Quinn *et al.*²⁰, suggests that such prolonged climate anomalies may recur from time to time.

As well as showing a decrease in intensity, the trade winds did not extend as far west across the ocean during the prolonged westerly anomaly. The weakening of the winds east of Australia show that the South Pacific Convergence Zone (SPCZ) was also reduced in intensity between 1940 and 1944. It should be noted that while this period saw the anomalous circulation prevailing over most of the Pacific, persistent local anomalies began as early as the mid-1930s in some regions such as the eastern equatorial Pacific (see Fig. 1). The prolonged westerly anomaly collapsed dramatically in 1946.

Fig. 3 *a*, Annual average wind stress for 1920–1983. A circle represents a mean wind of under 0.5 m s^{-1} . *b*, Wind anomaly during the period 1940–1944 compared to the climatology of Fig. 3*a*. Note that the scale is double that in (*a*). Only differences significant at the 95% level are plotted. Insignificant differences are represented by a circle, as are all differences less than 0.5 m s^{-1} . Locations with less than 60 observations are not marked.

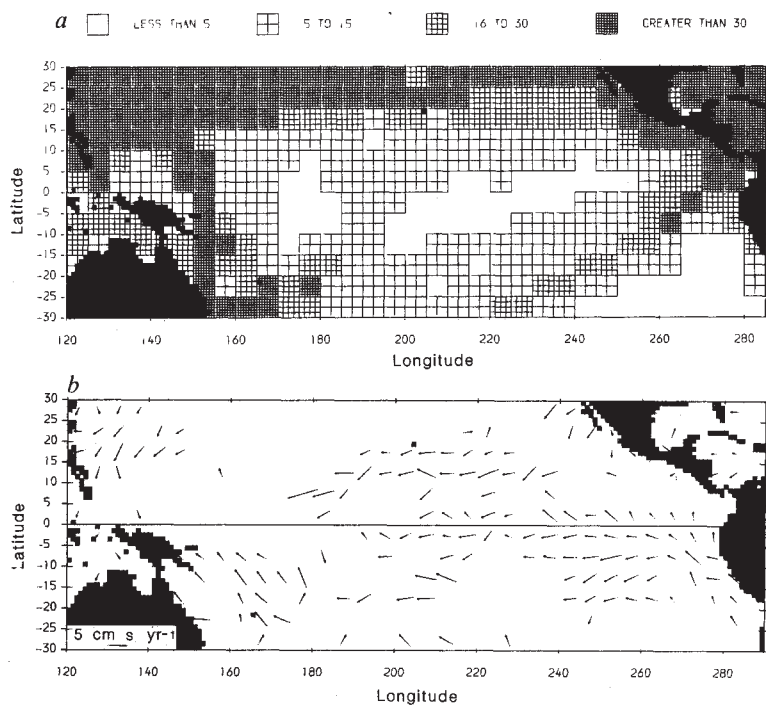
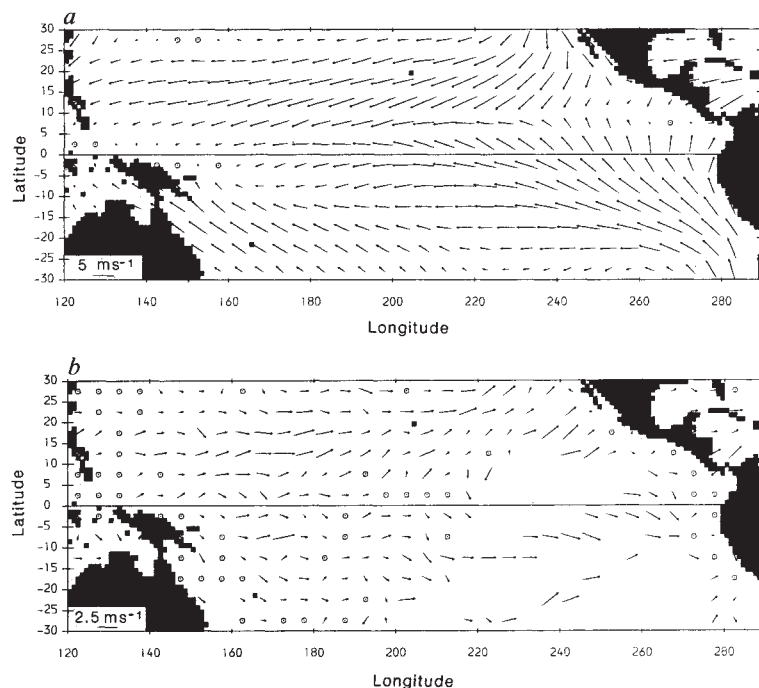


Fig. 4 *a*, Monthly mean number of surface wind observations for the period 1950–1983. *b*, Wind trends over the period 1950–1981. The vectors comprise the zonal and meridional trends. Only differences significant at the 95% level are plotted.

There is obviously the possibility that these observed extremes are due to poor data coverage during the Second World War. However in some regions of the Pacific the observational density was greater in this period than at other times. Time series similar to Fig. 2 for such regions (not shown here) still exhibit the climate anomaly. It should also be noted that the Quinn *et al.*²⁰ catalogue of El Niños includes the years 1939–1941, 1943, 1944 and 1946 as warm events. We believe that the whole period of 1939 to 1946 should be viewed as a major climate anomaly rather than as a succession of El Niños.

Another feature clearly visible in Fig. 2 is the trend towards more easterly winds as the present day is approached. In the South Pacific there is a gradual increase, while in the North the effect was more of a jump in the late 1940s. To quantify these trends a linear regression analysis was performed at every 5° by 5° square for the period 1950–81. In those cases when a square

had no data for a particular month climatology was used, though as data coverage for the period is relatively high (see Fig. 4*a*), this was infrequent. Those trends significant at the 95% level are plotted as vectors in Fig. 4*b*, and it can be seen that over much of the eastern Pacific there was an intensification in circulation of over 3 cm s^{-1} per year. This represents an increase in the wind between 1950 and 1981 of 1 m s^{-1} , which is about 20% of the climatological mean. Also clearly visible is an intensification of the meridional flow in the western Pacific, with the development of a strong southerly component north of the equator which is not apparent in the 64-year climatology of Fig. 3*a*. Figure 4*b* demonstrates that there has been a large change in the tropical Pacific surface wind climatology over the past 30 years.

The combined zonal and meridional strengthening of the trades since 1950 suggests an intensification of the Pacific cell

of the Hadley circulation and possibly a stronger Intertropical Convergence Zone (ITCZ). This is consistent with the observed cooling of the atmosphere between 1945 and 1970^{1,2,5-8} and the slight warming of the ocean during the same period^{2,9,10}. These would combine to give increased evaporation and therefore stronger convection in the ITCZ. It is of interest that there is evidence in our data for the trend to decrease and even reverse in the mid-1970s (see Fig. 2a). This is concomitant with the epoch at which some authors have suggested that the atmosphere began warming again^{1,4-7}.

The strengthening of the trade winds since 1950 has not been restricted to the Pacific, but also occurred in the Atlantic trade winds^{3,15,16}. Bunker¹⁶ found that in the period 1948–1972 the trade winds in the Atlantic strengthened by up to 4 cm s⁻¹ per year, a trend significant at the 95% level. It has been suggested²¹ that alterations in the trade-wind system will cause changes in tropical precipitation, in the location and intensity of the ITCZ and, perhaps, therefore has caused the dramatic decrease in Sahel rainfall since its maximum in 1950²². Thus, the observed strengthening of the winds reported here may be just one part of a widespread trend in the tropical climate.

We would like to thank Pat Roberts and Dave Ireland of the UK Meteorological Office for preparing the Pacific wind data

set used in this study, and Keith Shine for useful comments on the manuscript. This research was supported by the NERC.

Received 22 December 1986; accepted 26 March 1987.

1. Folland, C. K., Parker, D. E. & Kates, F. E. *Nature* **310**, 670–673 (1984).
2. Barnett, T. P. *Mon. Weath. Rev.* **112**, 303–312 (1984).
3. Arfl, R. *Can. J. Fish. aquat. Sci.* **42**, 1969–1978 (1985).
4. Jones, P. D., Wigley, T. M. L. & Wright, P. B. *Nature* **322**, 430–434 (1986).
5. Jones, P. D. *et al. J. Clim. appl. Met.* **25**, 161–179 (1986).
6. Jones, P. D., Wigley, T. M. L. & Kelly, P. M. *Mon. Weath. Rev.* **110**, 59–72 (1982).
7. Vinnikoy, K. Ya *et al. Meteor. Gidrol.* **6**, 5–17 (1980).
8. Saur, J. F. T. *J. appl. Met.* **2**, 417–425 (1963).
9. Paltridge, G. W. & Woodruff, S. *Mon. Weath. Rev.* **109**, 2427–2434 (1981).
10. Paltridge, G. W. *Mon. Weath. Rev.* **112**, 1093–1095 (1984).
11. Mitchell, J. M. Jr *Ann. N.Y. Acad. Sci.* **95**, 235–250 (1961).
12. Rasmussen, E. M. & Carpenter, T. H. *Mon. Weath. Rev.* **110**, 354–384 (1982).
13. Gill, A. E. & Rasmussen, E. M. *Nature* **306**, 229–234 (1983).
14. Philander, S. G. *Nature* **302**, 295–301 (1983).
15. Servain, J., Picaut, J. & Bussalachi, A. J. in *Coupled Ocean–Atmosphere Models* (ed. Nihoul, J.) 211–237 (Elsevier, Amsterdam, 1985).
16. Bunker, A. F. *Mon. Weath. Rev.* **108**, 720–732 (1980).
17. Goerss, J. S. & Duchon, C. E. *J. phys. Oceanogr.* **10**, 478–479 (1980).
18. Barnett, T. P. *J. Atmos. Sci.* **34**, 221–236 (1977).
19. Hess, S. L. *Introduction to Theoretical Meteorology* (Holt, Reinhart and Winston, New York, 1959).
20. Quinn, W. H., Zopf, D. O., Short, K. S. & Yang, R. T. W. *Fishery Res. Bull.* **76**, 663–678 (1978).
21. Ellsaesser, H. W., MacCracken, M. C., Walton, J. J. & Grotch, S. L. *Rev. Geophysics* **24**, 745–792 (1986).
22. Folland, C. K., Palmer, T. N. & Parker, D. E. *Nature* **320**, 602–607 (1986).

Higher-order eccentricity cycles of the middle and late Miocene climatic variations

R. K. Tiwari*

Theoretical Geophysics Group, National Geophysical Research Institute, Uppal Road, Hyderabad 500007 AP, India

Although there is evidence for Milankovitch cycles in deep-sea sediments from the Pleistocene^{1–4} and Upper Miocene⁵, the origin of orbital eccentricity cycles in these records is controversial^{5–14}. Recent spectral analyses^{15,16} of non-glacial climatic records of mid-Cretaceous and early Mesozoic exhibit the dominance of 100- and 400-kyr eccentricity cycles and exclude the possibility that the above periodicities originate from the internal variability of ice sheets. Here I present the Walsh spectrum analyses of 10 Myr oxygen ($\delta^{18}\text{O}$) and carbonate ($\delta^{13}\text{C}$) isotope variations, spanning the middle and late Miocene, which reveal statistically significant periodicities of approximately 2.0 and 1.25 Myr, 800, 400, 115 and 93 kyr and harmonics of the higher-order terms with remarkable resolution. Higher-order periodicities, hitherto unexplained in the spectrum of deep-sea records, match quite well with predicted large eccentricity cycles in the Earth's orbital variations¹⁷, indicating that eccentricity forcing (induced by orbital variations) exerts a major influence on the Earth's climate.

The global climate underwent major change during the Miocene period. The data indicate that there were several significant fluctuations at regular intervals of approximately 1–1.3 Myr. According to Kennet¹⁸ episodes interpreted as having been warmer and with less polar ice occurred in 13.5–4 Myr intervals, between about 13–12 Myr (late-middle Miocene), 7.5–6.2 Myr (middle-late Miocene) and 5.2–4 Myr (early Pliocene). Episodes interpreted as having been cooler and with more ice occurred from about 11–10 Myr (earliest late Miocene) and 6.2–5.2 Myr (latest Miocene). Spectral analysis suggests that the Upper Miocene climatic history significantly differs from the Pleistocene and reveals its sensitivity to low-frequency variations⁵. The questions I shall examine here are (1) evidence for

Table 1 Data for the core from site 588

Locations and other details	Sedimentation rates	Carbonate accumulation rates
Latitude 26°06.70' S Longitude 161°13.60' E Water depth 1,533 m Modern water mass Warm-subtropical Thickness of Miocene sequence drilled 288 m Depth of ooze/chalk transition 260 m	1.5 to 3 cm × 10 ³ throughout the Miocene	Much more variable from ~15 to 10 Myr (middle Miocene) reaching a peak (up to 13 g cm ⁻² per 10 ³ yr) at ~12 Myr. Between 10 and 14 Myr fairly constant and average rates of accumulation except for slightly higher rates at 5 Myr near the Miocene/Pliocene boundary.

Data from ref. 18.

higher-order periodicity in a period range of 100 kyr to 2,000 kyr using a modern spectral technique in 10-Myr-long Miocene records of oxygen isotope and carbonate isotope variations; (2) its correlation with predicted large eccentricity cycles in orbital variations; and (3) the probable mechanism that drove Miocene climatic variability.

The isotope data under study come from a part of a 315-m thick core at site 518 in the south Pacific under the Deep Sea Drilling Project (DSDP) Leg programme. A detailed account of the data characteristics is given elsewhere¹⁸. However, a brief summary of the data relevant to this study is presented in Table 1 and described here. The analysed core spans the middle to late Miocene epoch covering a period of about 10 Myr and contains relatively uniform light-coloured foraminifers bearing to foraminifers-rich nannofossil oozes to chalk throughout the Miocene. The calcium carbonate content is greater than 50% throughout and the terrigenous sedimentary components are virtually non-existent. Preservation of foraminiferal assemblages are generally excellent and deposition always occurred at a depth shallower than the foraminiferal lysocline. This sequence contains continuous, largely uninterrupted and high-resolution stratigraphic records which provide an excellent opportunity to study the Miocene climatic history. Rates of sedimentation,

* Present address: Institut für Geophysik, Olshausenstrasse 40-60, Christian Albrechts Universität, 23 Kiel, FRG.

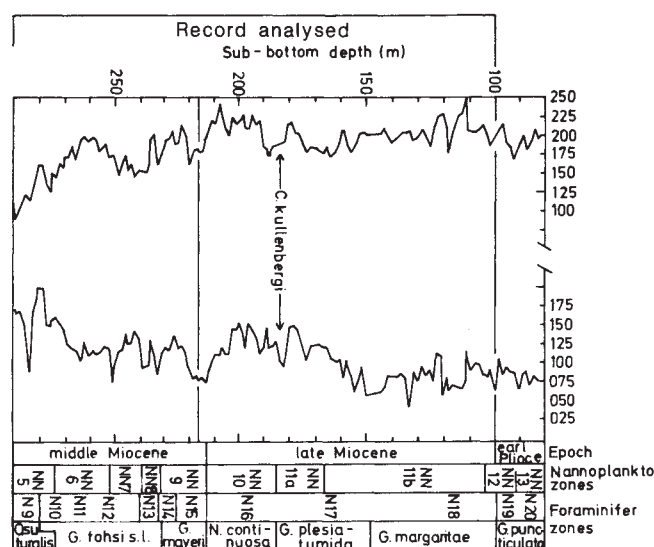


Fig. 1 Benthic $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ variations plotted against depth at site 588 for middle and late Miocene. Biostratigraphic zonation for calcareous nannoplankton and planktonic foraminifers. Data from Kennet¹⁸.

however, are variable throughout the Miocene period and therefore both records may lack detail.

For much of the Miocene, sedimentation rates ranged from 1.5 to 3 cm per 10^3 yr which gives a sampling interval of 25×10^3 yr¹⁸. Note that sampling intervals for isotope analysis are smaller than those used to define the positions of the calcareous nannoplankton boundaries and hence the isotope record is of higher resolution than the chronological framework¹⁸. It is prudent, therefore, to study higher-order periodicity in these records which may throw some light on long-term behaviour of the Miocene climate.

The chronological framework is based on ages assigned to Neogene calcareous nannoplankton¹⁸ zonal boundaries by Berggren¹⁹. Some uncertainty exists regarding the age of the middle/late Miocene boundary. Based on a correlation between chron 9 and anomaly 5, an age of 11 Myr has been assigned to the middle/late Miocene boundary. This has been modified based on the correlation between chron 11 and anomaly 5. Accordingly, an age of 8.92 to 10.42 Myr was assigned to chron 11 in accordance with the palaeomagnetic timescale of Berggren *et al.*²⁰. This puts the middle/late Miocene at an age of 10.42 Myr²⁰. Here, in order to avoid subjective bias, I have used the same timescale as Kennet¹⁸. Records of both carbon and oxygen isotopes exhibit discontinuous and non-sinusoidal

Table 2 Significant periodicities in 10 Myr long oxygen ($\delta^{18}\text{O}$) and carbonate ($\delta^{13}\text{C}$) isotope variations during the middle and late Miocene period: high resolution Walsh spectra

$\delta^{13}\text{C}$	$\delta^{18}\text{O}$	Predicted cycles of orbital eccentricity*
2.0 Myr	2.5 Myr	2.035 Myr
1.0 Myr	1.25 Myr	1.30 Myr
714 kyr	833 kyr	413×2 kyr
555 kyr	500 kyr	603 kyr
357 kyr	357 kyr	95×4 kyr
250 kyr	250 kyr	112×2 kyr
110 kyr	112 kyr	107 kyr
91 kyr	93 kyr	95 kyr

* From ref. 17.

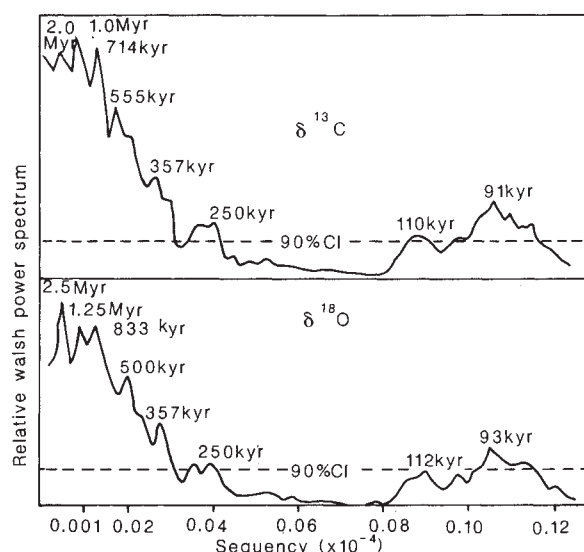


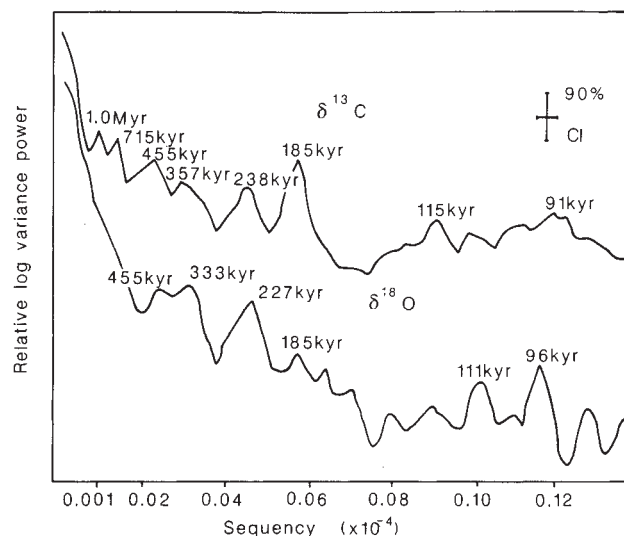
Fig. 2 High-resolution Walsh power spectra of carbonate isotope variations (top) and oxygen isotope variations (bottom). Walsh power spectra are smoothed using three points (0.25, 0.5, 0.25) Hanning window. Length of time series is 10 Myr covering the middle and late Miocene periods. CI indicates confidence intervals.

Table 3 Highly stable peaks in the average variance Walsh spectra

1.0 Myr	—	1.3 Myr
714 kyr	—	413×2 kyr
454,545 yr	454,545 yr	413,000 yr
357,143 yr	333,333 yr	$95,000 \times 4$ yr
238,000 yr	227,000 yr	$112,000 \times 2$ yr
185,000 yr	185,000 yr	$95,000 \times 2$ yr
115,000 yr	111,000 yr	107,000 yr
91,000 yr	96,000 yr	95,000 yr

patterns with alternating sharp peaks (Fig. 1). The Walsh transform technique is suitable for analysing periodicity from such signals decoded from deep-sea stratigraphic sequences. The appropriateness and resolving capability of the Walsh technique have been well emphasized and demonstrated on geomagnetic, palaeomagnetic, palaeoclimatic and several other geophysical time series^{4,21-23}. Before performing spectral analysis, both records were properly detrended and the resulting time series were then interpolated at uniform spacing of 7,812 yr. Initially, the Walsh transform technique was applied to the entire digital time series and the power spectra were computed using a periodogram approach. The computed power spectra were smoothed using a three-point Hanning window and a statistical test was performed to evaluate confidence limits for the smoothed spectral components which are depicted by dashed lines in Fig. 2. Records of both oxygen and carbon isotopes exhibit striking similarities in their spectral structure. These results are summarized in Tables 2 and 3 and compared with predicted higher-order eccentricity cycles¹⁷. It is clear from Table 2 that predicted eccentricity cycles show a close resemblance to climatic periodicities. Bearing in mind the intrinsic non-stationary properties of deep-sea records, as a consequence of which deceptive appearances of power peaks can result, the digital time series was split into two smaller sub-series and a similar analysis was performed. An average Walsh power spectrum of two sub-time series are displayed in Fig. 3. Figure 3 and Table 3 clearly show the persistence of the smaller periodicities observed in Fig. 2. One intriguing observation is the dominance of higher-order peaks (1 Myr and 714 kyr) in the carbonate spectra. Further, it is interesting to note that the double 93 and 115 kyr peak con-

Fig. 3 The log variance average Walsh power spectra smoothed with Hanning window and averaged over for four degrees of freedom. Spectral peaks are highly stable but poorly resolved towards the low frequency end of the spectrum. Concentration of strong power towards the low frequency end indicate the dominance of low frequency orbital eccentricity forcing during the middle and late Miocene periods.



sistently appeared in all the above computations instead of one single peak at 100 kyr. The two periodicities are related to a beat frequency arising from the splitting of precessional cycles¹⁷. Present analysis reveals, for the first time, almost all the periodicities in deep-sea geological records as predicted in eccentricity variations of the Earth's orbit¹⁷.

Discovery of higher-order eccentricity cycles in deep-sea geological records suggests a possible link between the two phenomena. Although calculated eccentricity variations should not have a significant effect on mean annual solar radiation²⁴, evidence for eccentricity cycles as a dominating spectral component has been invariably reported in several climatic indicator records^{1,2,5,15,16,25-29}. Thus, the origin of eccentricity cycles in these records is controversial⁵⁻¹⁴. Several reports⁷⁻¹⁰ have ascribed eccentricity cycles to nonlinearity associated with ice sheet growth, decay and internal variability whereas others^{7,11,14} have suggested that eccentricity peaks in the climatic records may result from the climatic system's nonlinear response to the precessional signal. Alternatively, eccentricity forcing may exert a major direct control on the Earth's climate^{2,15,16}. Cross-spectral analysis of the upper Miocene calcium carbonate records and orbital eccentricity variations exhibits high coherence at 400- and 100-kyr eccentricity peaks⁵ and support the latter view. Most recent spectral analyses of non-glacial climatic records of mid-Cretaceous and early Mesozoic eliminate the possibility that the 400- and 100-kyr cycles are the result of the internal variability of ice sheets and convincingly reinforce the impression that the noted cycles are primarily of astronomical origin^{15,16}. Because of the stability and dominance of higher-order periodicities in the two climatic records and their striking resemblance with predicted eccentricity cycles I agree with the above view. Although the eccentricity effect on annual mean solar radiation is weak, recent theoretical climatic models^{30,31} demonstrate the

possibility of amplification of the eccentricity signal by the climatic system itself. According to these models the external periodic forcing due to orbital variations combined with an internal stochastic mechanism amplifies the small external forcing which, in turn, induces a large amplitude with long-term alterations of the climate.

Besides providing support for a debated scientific issue, this work also suggests that higher-order periodicities in the climatic records are also remarkably correlated with eccentricity cycles. Slight departure from ideal periodicity may arise due to variable sedimentation rates as previously noted. Due to a lack of detail, however, it is not appropriate at this stage to assign any exact numerical value to such errors at different frequencies. It is unlikely, however, that such errors will finally influence our conclusion, in any case, on this timescale. Further investigations using more refined data from areas of possibly uniform sedimentations and climatically sensitive zones, may perhaps elicit better pictures. Although one may argue that higher-order eccentricity cycles may also arise due to several other factors^{8-11,13,14}, to the best of my knowledge however, there exists no published or calculated evidence to support such arguments.

In view of the recent findings of eccentricity cycles in non-glacial climate records and the results here, I conclude that orbital eccentricity forcing is one of the most plausible candidates that may have a major influence on the terrestrial climate.

I thank Professor R. Meissner and Professor J. Zschau for their support and valuable suggestions. Thanks are also due to K. Helbig, G. Reim and L. Bittner for their help in preparing the manuscript. I thank Dr J. G. Negi, Deputy Director of NGRI Hyderabad who has always been a source of inspiration for my research work. This work has been carried out under a Deutscher Akademischer Austauschdienst exchange programme at the University of Kiel, FRG.

Received 17 November 1986; accepted 24 February 1987.

- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).
- Imbrie, J. et al. in *Milankovitch and Climate* part 1 (eds Berger, A., Imbrie, J., Hays, J., Kukla, G. & Salzman, B.) 269-305 (Reidel, Boston, 1984).
- Berger, A. L. *Celest. Mech.* **15**, 53-74 (1975).
- Negi, J. G. & Tiwari, R. K. *Earth planet. Sci. Lett.* **70**, 138-148 (1984).
- Moore, T. C., Pisias, N. G. & Dunn, D. A. *Mar. Geol.* **46**, 217-233 (1982).
- Shackleton, N. J. & Opdyke, N. D. *Mem. geol. Soc. Am.* **145**, 499-464 (1973).
- Le Treut, H. & Ghil, M. *J. geophys. Res.* **88**, 5167-5190 (1983).
- Pollard, D. P. *Nature* **296**, 334-338 (1982).
- Weertman, J. *Nature* **261**, 17-20 (1976).
- Birchfield, G. E. *J. geophys. Res.* **82**, 4909-4913 (1977).
- Imbrie, J. & Imbrie, J. Z. *Science* **194**, 943-952 (1980).
- Oerlemans, J. *Nature* **287**, 430-432 (1980).
- Kukla, G. J., Berger, A., Lotti, R. & Brown, J. *Nature* **290**, 295-300 (1981).
- Wigley, T. M. L. *Nature* **264**, 629-631 (1981).
- Herbert, T. D. & Fischer, A. G. *Nature* **321**, 739-743 (1986).

- Olsen, P. E. *Science* **234**, 842-848 (1986).
- Berger, A. L. *Nature* **269**, 44-45 (1977).
- Kennet, J. P. *Init. Rep. DSDP XC, Part 2*, 1383-1420 (1986).
- Berggren, W. A. *Proc. IGCP-114 Int. Workshop Pacific Neogene Biostratigraphy* Nov. 25-29, Osaka, Japan 29-60 (1981).
- Berggren, W. A., Kent, D. V., Flynn, J. J. & Van Couvering, J. A. *Geol. Soc. Am. Bull.* **96**, 1407-1412 (1985).
- Negi, J. G. & Tiwari, R. K. *Geophys. Res. Lett.* **10**, 313-316 (1983).
- Tiwari, R. K. thesis B.H.U. Varanasi, India (1985).
- Weedon, G. P. *Earth planet. Sci. Lett.* **76**, 321-335 (1986).
- Wollin, G., Ryan, W. B. F. & Ericson, D. B. *Earth planet. Sci. Lett.* **41**, 395-397 (1978).
- Briskin, M. & Harrel, J. *Mar. Geol.* **36**, 217-233 (1982).
- Berger, W. H. *J. foram. Res.* **3**, 187-195 (1973).
- Arrhenius, G. *Rep. Swed. Deep Sea Exped.* **1**, 1-228 (1952).
- Hays, J. D., Saito, T., Opdyke, N. D. & Bruckle, L. H. *Geol. Soc. Am. Bull.* **80**, 1481-1514 (1969).
- Broecker, W. S. & Von Donk, J. *Rev. Geophys. Space Phys.* **8**, 169-198 (1968).
- Nicolas, C. *Tellus* **34**, 1-9 (1982).
- Benzi, R., Parisi, G., Suter, A. & Vulpiani, A. *Tellus* **34**, 10-16 (1982).

Temperatures of evaporating clusters

Cornelius E. Klotz

Chemical Physics Section, Health and Safety Research Division,
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

The fabrication of new forms of materials through sudden changes in their temperatures is a familiar, if mysterious, process. Thus, it is not known why some substances can be quenched to yield glasses, while others 'prefer' to crystallize. The appearance of quasicrystals in rapidly cooled metallic alloys further illustrates the surprises and challenges that continue to appear. One way to obtain rapid cooling is to isolate a small aggregate of matter: cooling then results as a consequence of the ensuing evaporation. The temperature attained in this process is apt, more than any other parameter, to provide an incisive index to other properties of interest. Here I show that these temperatures, as well as their property of scaling with Lennard-Jones parameters, may often be understood in terms of the 'evaporative ensemble'; it is required only that evaporation be occurring.

Isolated aggregates of matter can be prepared by a number of techniques. Temperatures have been reported¹⁻⁴ for clusters generated in supersonic expansions, and so we shall begin with a discussion of them. In this method a gas is permitted to expand through a small pinhole into a high vacuum. Under appropriate conditions the quasi-isentropic expansion may result in massive condensation. As the density drops, however, collisions cease. The clusters are then investigated¹⁻⁴ at a point further downstream with an electron diffraction technique.

The evaporative ensemble^{5,6} assumes that during this period of free flight each aggregate has suffered at least one evaporation. In common with more familiar ensembles, all other information is eschewed. The state of a given cluster is then characterized solely by the condition that its first-order rate constant for further evaporation is $\sim t^{-1}$, where t is the time-lapse since the initial stage of condensation.

Explicit calculations^{5,6} of these rate constants have suggested that a modified version of Trouton's relation⁷ will be applicable,

$$RT/\Delta E = \text{constant} \quad (1)$$

where R is the gas constant and ΔE the molar energy of evaporation. The 'constant' on the right-hand side of equation (1) is understood to be roughly independent of the material constituting the cluster; it is necessarily a decreasing function of time (as the cluster cools off), but only weakly so. On a timescale of tens of microseconds, the calculations suggest a typical value $\approx 4 \times 10^{-2}$.

I have used this value to predict the temperatures of the clusters studied by the electron diffraction technique. The thermal data required by equation (1) were taken from standard sources^{8,9}. A comparison with the experimental data is shown

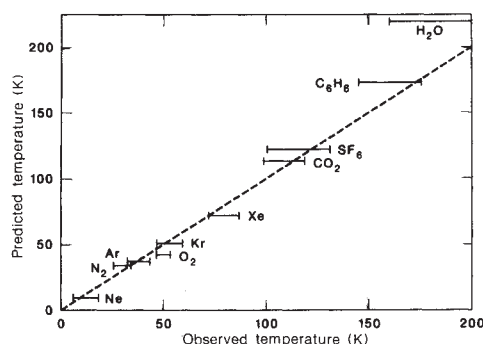


Fig. 1 Predicted and observed (see refs 1-4) temperatures of clusters.

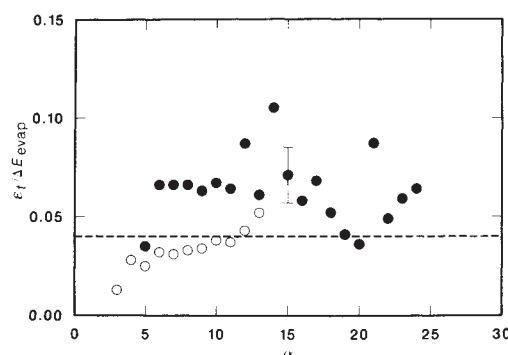


Fig. 2 Scaled kinetic energy release in cluster evaporation, from refs 12 and 13. The dashed line indicates the prediction of equation (1). A typical error bar is illustrated. ●, argon; ○, CO₂.

in Fig. 1. The line has unit slope and is not necessarily a 'best-fit.'

Other authors¹ have considered a scaling between the cluster temperature and the Lennard-Jones energy parameter. The relation implied by equation (1) is clearly similar, and a simple proportionality between temperature and energy-of-evaporation would have added nothing new. The quantitative character of the correlation observed here, however, directly demonstrates that these clusters may be understood in terms of the evaporative ensemble. The scaling of their properties may be accordingly rationalized.

We remark that this ensemble was defined originally^{5,6} in a discussion of 'metastable evaporation' and that it was in that context that the parameters on the left-hand side of equation (1) first arose. If an ensemble of clusters containing α monomeric units is isolated after time t , its population P_α will decay according to the law

$$-\delta P_\alpha \approx P_\alpha (C/R) (RT/\Delta E)^2 \cdot \delta \ln t \quad (2)$$

for large clusters and in the absence of magic-number effects. C , the heat capacity per monomeric unit, scarcely changes among elementary materials. The correlation in Fig. 1 is therefore linked with the existence of an important uniformity in the phenomenon of metastable evaporation.

Gspann¹⁰ has also suggested that these temperatures may be understood in terms of the kinetics of evaporation. The expression which he used to examine this contains some ambiguities, whereas equations (1) and (2) hinge on little more than hard-sphere collision theory and microscopic reversibility. They thus provide a rather more convincing test of Gspann's idea. It should be noted that an evaporative ensemble cannot strictly be characterized by a single, well-defined temperature. For clusters sufficiently large as to yield electron diffraction patterns diagnostic of bulk properties, this is not important. For much smaller clusters, however, a 'temperature' is to be understood only as signifying an energy—in particular the high-energy edge of the ensemble, and that from which evaporation is occurring.

To accommodate the smaller aggregates I turn to a second and more general sort of thermometer. The flux distribution of kinetic energies in the reaction coordinate of a dissociating species is predicted¹¹ by quasi-equilibrium theory to be that of a two-dimensional Maxwell-Boltzmann distribution. The average such energy $\bar{\epsilon}^\ddagger = kT^\ddagger$. If this distribution persists upon further separation of the fragments, then $kT^\ddagger$ may be identified with $\bar{\epsilon}$, the measurable kinetic energy release. It will also be less than the 'temperature' of the parent molecule, because of the cooling which accompanies the fragmentation.

Measurements of kinetic energy release in evaporating systems have been reported^{12,13}, as have calculations of their magnitude^{13,14}. These latter made use of standard methods and thus did not take advantage of well-known similarities among liquids, in particular those embodied in Trouton's rule. I therefore compare in Fig. 2 the measurements of $\bar{\epsilon}$ with the kT defined by equation (2). It seems clear that, in spite of the several

qualifications noted above, the temperature scales defined by the disparate phenomena of kinetic energy release and electron scattering are related; with comparably-sized clusters this relation would presumably be even more evident.

Other techniques, such as laser ablation and sputtering, have also been used to prepare isolated aggregates of matter. The metastability of sputtered copper clusters¹⁵ has already been shown⁶ to be consistent with the evaporative ensemble. We may thus suppose that the scaling described here will be widely applicable. It is required only that each cluster will have undergone at least one evaporation.

This condition is not always met. Valente and Bartell^{3,4} have noted that lower temperatures may be obtained from supersonic expansions when the condensing gas is diluted in a non-condensing carrier. The kinetics of evaporation thus establish only an upper limit to the range of accessible temperatures. Considerable flexibility in the tailoring of the magnetic and other properties of the aggregates remains. We observe, for example, that the Curie points of the common ferromagnetic materials lie below the demarcations given by equation (1).

I thank G. Torchet and A. J. Stace for stimulating communications. This research was sponsored by the Office of Health and Environmental Research, US Department of Energy under contract with Martin Marietta Energy Systems, Inc.

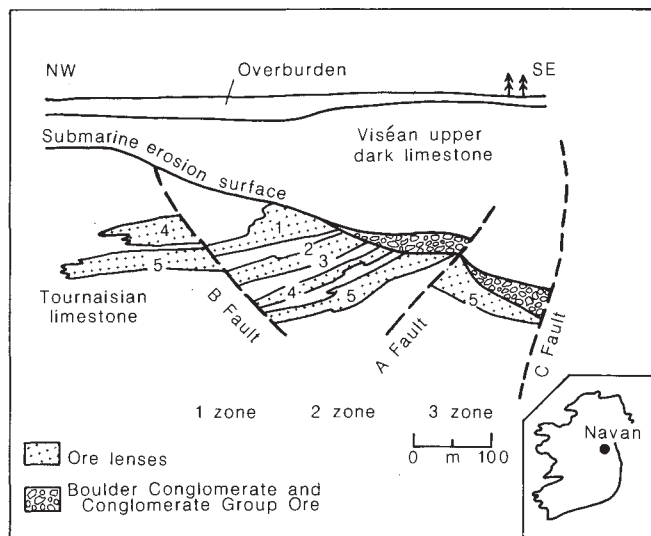


Fig. 1 Generalized cross-section of the Navan orebody showing the lens structure referred to in the text (after Andrew and Ashton²⁹). Inset: map of Ireland showing location of Navan.

Received 23 December 1986; accepted 25 March 1987.

1. Farges, J., de Feraudy, M. F., Raoult, B. & Torchet, G. *Surf. Sci.* **106**, 95-100 (1981).
2. Torchet, G., Schwartz, P., Farges, J., de Feraudy, M. F. & Raoult, B. *J. chem. Phys.* **79**, 6196-6202 (1983).
3. Valente, E. J. & Bartell, L. S. *J. chem. Phys.* **79**, 2683-2686 (1983).
4. Valente, E. J. & Bartell, L. S. *J. chem. Phys.* **80**, 1451-1457 (1984).
5. Klots, C. E. *J. chem. Phys.* **83**, 5854-5860 (1985).
6. Klots, C. E. *Z. Physik D38*, 5 (1987).
7. Trouton, F. *Phil. Mag.* **18**, 54-56 (1884).
8. Huttman, R. et al. (eds) *Selected Values of the Thermodynamic Properties of the Elements* (Am. Soc. Metals, Metals Park, Ohio, 1973).
9. Raznjevic, K. *Handbook of Thermodynamic Tables and Charts* (Hemisphere Publishers, Washington, 1976).
10. Gspann, J. *Physics of Electronic and Atomic Collisions* (ed. Datz, S.) 79-96 (North Holland, Amsterdam, 1982).
11. Klots, C. E. *J. chem. Phys.* **58**, 5364-5367 (1973).
12. Stace, A. J. & Stukla, A. K. *Chem. Phys. Lett.* **85**, 157-160 (1982).
13. Stace, A. J. *J. chem. Phys.* **85**, 5774-5778 (1986).
14. Illies, A. J., Jarrold, M. F., Bass, L. M. & Bowers, M. T. *J. Am. chem. Soc.* **105**, 5775-5781 (1983).
15. Begeman, W., Meiwes-Broer, K. H. & Lutz, H. O. *Phys. Rev. Lett.* **56**, 2248-2252 (1986).

Origin of a giant orebody at Navan, Ireland

H. Mills*†, A. N. Halliday†§, J. H. Ashton‡, I. K. Anderson* & M. J. Russell*

* Department of Applied Geology, University of Strathclyde, Glasgow G1 1XJ, UK

† Scottish Universities Research & Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

‡ Tara Mines, Navan, Ireland

The Navan base-metal deposit in Ireland formed 360 Myr ago¹ and is one of the world's largest orebodies, with 70 million tonnes of 12% Zn+Pb. Much of the mineralization was sedimentary, exhalative: other well known examples of this class of the so-called SEDEX deposits are Mount Isa in Australia, Sullivan in Canada, and Meggen in Germany. Several models for the formation of such deposits have been suggested. Here we report the first systematic stratigraphic Pb isotopic study of galenas in any giant exhalative Zn+Pb orebody. We document evidence that the sources of Pb in the Navan orebody evolved to produce increased dispersion and less radiogenic averages of lead isotopic compositions. This

is as expected if the orebody formed from a hydrothermal convection cell which expanded to scavenge metals from progressively deeper parts of the continental crust.

Various genetic models have been suggested for the origin of SEDEX base-metal deposits: (1) basin brine expulsion, where saline connate waters dissolve metals from compacting sediments and are expelled upwards during basin subsidence²⁻⁶; (2) a hydrothermal convection system where fluid circulation is initiated and added to by intrusive igneous activity⁷⁻¹⁰; (3) release of metal-bearing waters during metamorphism¹¹⁻¹³; (4) seismic pumping in which fluids fill cracks which have opened up before a shear failure event which then expels the (heated) fluids back upwards¹⁴; (5) downward excavating convection of modified heated saline metalliferous sea water in which, during a period of crustal extension, the fluids penetrate as far as the brittle/ductile transition zone which is advanced to progressively greater depths by the concomitant cooling^{15,16}.

Distinguishing between these models using Pb isotopic compositions is partly based upon observations for other deposits and partly upon expectations derived from the process involved. Basin brine expulsion (model 1) has been most frequently invoked to explain Mississippi Valley type (MVT) deposits. These are commonly found to contain very radiogenic and exceedingly variable Pb isotopic compositions^{17,18}. Deposits of magmatic affiliation (model 2) should have relatively homogeneous Pb isotopic compositions which are similar to those of contemporaneous magmas. Deposits formed by extraction of metamorphic fluids (model 3) obviously would have no specific Pb isotope signature because this would depend simply on what is being metamorphosed¹⁹. We doubt that model 3 is relevant to giant-sized SEDEX deposits, because volumes of metamorphic fluid would probably be relatively small and the major tectonic thickening and metamorphism in this particular region had ceased 40 Myr previously (400 Myr ago). Nonetheless the possibility that a metamorphic basement was involved can obviously be investigated using Pb isotopes provided it has had a different history than that of the overlying crust. Seismic pumping (model 4) has never been invoked as a dominant mechanism in the formation of SEDEX deposits because it is difficult to see how the process can be sufficiently long-lived and powerful to produce major orebodies. It has no specific predictions in terms of lead isotope results. The model of deepening convective systems (model 5) may be tested with a lead isotope experiment. In general terms, high-grade metamorphic

§ Present address: Department of Geological Sciences, University of Michigan, Ann Arbor Michigan 48109-1063, USA.

Table 1 Lead isotope analyses of galenas from Navan

Location	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$	Location	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
Lens 1				269.5 m			
SH652	18.015 $\pm$ 3	15.522 $\pm$ 4	37.767 $\pm$ 7	SH652	18.070 $\pm$ 18	15.564 $\pm$ 26	38.023 $\pm$ 30
220.5 m				275 m	18.100 $\pm$ 6	15.593 $\pm$ 15	38.117 $\pm$ 20
SH652	18.090 $\pm$ 10	15.540 $\pm$ 8	37.886 $\pm$ 20	Lens 4			
220.5 m	18.105 $\pm$ 8	15.549 $\pm$ 7	37.920 $\pm$ 13	SH652	18.090 $\pm$ 6	15.500 $\pm$ 9	37.584 $\pm$ 17
SH652	17.551 $\pm$ 5	15.419 $\pm$ 4	37.198 $\pm$ 10	297.5 m			
227.5 m	17.560 $\pm$ 10	15.425 $\pm$ 8	37.203 $\pm$ 12	UG226 N	18.185 $\pm$ 11	15.575 $\pm$ 15	38.075 $\pm$ 27
SH652	17.858 $\pm$ 10	15.532 $\pm$ 5	37.556 $\pm$ 13	1,390 Level			
227.5 m				UG226 N	18.115 $\pm$ 3	15.580 $\pm$ 2	37.970 $\pm$ 7
SH652	17.841 $\pm$ 12	15.525 $\pm$ 11	37.664 $\pm$ 21	1,390 Level			
232.5 m				UG253 access	18.215 $\pm$ 8	15.553 $\pm$ 10	38.022 $\pm$ 17
SH652	17.998 $\pm$ 10	15.507 $\pm$ 15	37.664 $\pm$ 30	1,315 Level			
241 m	18.005 $\pm$ 12	15.512 $\pm$ 17	37.680 $\pm$ 35	UG253 access	18.212 $\pm$ 10	15.585 $\pm$ 8	38.125 $\pm$ 16
SH652	17.791 $\pm$ 3	15.474 $\pm$ 3	37.431 $\pm$ 6	1,315 Level			
242.5 m				Lens 5			
SH652	18.202 $\pm$ 6	15.644 $\pm$ 10	38.386 $\pm$ 13	SH652	18.155 $\pm$ 4	15.601 $\pm$ 4	38.079 $\pm$ 7
246.5 m				325 m	18.175 $\pm$ 10	15.620 $\pm$ 10	38.093 $\pm$ 15
UG229 N	18.203 $\pm$ 3	15.561 $\pm$ 3	38.043 $\pm$ 10	SH652	18.222 $\pm$ 6	15.610 $\pm$ 5	38.191 $\pm$ 14
1,435 Level	18.191 $\pm$ 7	15.585 $\pm$ 10	38.152 $\pm$ 14	330.5 m			
UG229 N	18.201 $\pm$ 3	15.553 $\pm$ 2	38.011 $\pm$ 11	SH652	18.186 $\pm$ 15	15.542 $\pm$ 13	37.968 $\pm$ 22
1,435 Level				332.5 m			
UG229 N	18.204 $\pm$ 14	15.592 $\pm$ 13	38.162 $\pm$ 29	SH543	18.170 $\pm$ 21	15.554 $\pm$ 15	38.056 $\pm$ 31
1,435 Level				335 m			
SH543	17.910 $\pm$ 2	15.562 $\pm$ 3	37.764 $\pm$ 8	SH543	18.170 $\pm$ 16	15.563 $\pm$ 10	38.000 $\pm$ 21
224.5 m				338.5 m			
SH543	18.190 $\pm$ 19	15.553 $\pm$ 13	37.848 $\pm$ 27	SH543	18.207 $\pm$ 10	15.606 $\pm$ 7	38.220 $\pm$ 16
252.7 m				330 m			
Lens 2				UGBL6. P8	18.183 $\pm$ 15	15.550 $\pm$ 13	38.052 $\pm$ 21
(1,285 Level)				1,315 Level			
UG2217	18.197 $\pm$ 8	15.584 $\pm$ 5	38.129 $\pm$ 14	Silvermines			
UG6201	18.172 $\pm$ 5	15.553 $\pm$ 6	38.015 $\pm$ 11	SH26	18.161 $\pm$ 9	15.528 $\pm$ 3	37.933 $\pm$ 20
UG6202	18.174 $\pm$ 11	15.563 $\pm$ 9	37.998 $\pm$ 28	4520	18.196 $\pm$ 10	15.556 $\pm$ 10	38.053 $\pm$ 21
UG6203	18.022 $\pm$ 15	15.563 $\pm$ 9	37.796 $\pm$ 22	4700-3	18.173 $\pm$ 6	15.544 $\pm$ 6	37.982 $\pm$ 15
	18.040 $\pm$ 12	15.572 $\pm$ 8	37.821 $\pm$ 24	Tynagh			
UG6204	18.101 $\pm$ 19	15.557 $\pm$ 21	37.799 $\pm$ 31	CF123	18.028 $\pm$ 4	15.559 $\pm$ 2	37.943 $\pm$ 11
Lens 3				CF123	18.004 $\pm$ 10	15.529 $\pm$ 13	37.879 $\pm$ 15
SH652	18.216 $\pm$ 6	15.634 $\pm$ 4	38.335 $\pm$ 8		18.010 $\pm$ 9	15.535 $\pm$ 11	37.900 $\pm$ 12
268.5 m	18.209 $\pm$ 4	15.614 $\pm$ 10	38.320 $\pm$ 6	3/5000	18.026 $\pm$ 8	15.550 $\pm$ 6	37.939 $\pm$ 15
SH652	18.098 $\pm$ 16	15.546 $\pm$ 9	38.001 $\pm$ 25				

Samples dissolved in dilute HNO_3 , purified by electrodeposition³⁵, and analysed (using single Re filament with silica gel and phosphoric acid) on the V.G. Isomass 54E mass spectrometer at SURRC. Measurements were corrected for mass discrimination by normalizing to the recommended values for the NBS 981 standard. Errors refer to least significant digits and are 2σ mean run precision. Long-term reproducibility is estimated at closer to $\pm 0.1\%$. SH, surfacehole; UG, underground.

basement such as makes up the lower and middle crust in this area of Ireland²⁰, can be expected to have relatively unradiogenic Pb—the time-integrated effect of uranium depletion. Therefore a progression to less radiogenic values with time will be recorded in the lead isotope ratios of sulphides precipitated from the fluids which have extracted metals from progressively deeper levels in the crust.

The Navan orebody (Fig. 1), operated by Tara Mines Ltd, is one of several unmetamorphosed base-metal SEDEX deposits in Ireland which also includes Tynagh and Silvermines^{20–27}. Navan was chosen because of its size and because of the clear stratigraphic superposition of the lenses^{28,29}. There are five such lenses, numbers 2 and 3, usually classed as one unit (Fig. 1). The main Pb+Zn mineralization at Navan is hosted by Tour-naianian bioclastic and oolitic limestones, micrites and minor sandstones. This sequence unconformably overlies Lower Palaeozoic shales and grits with Ordovician volcanic horizons. A syenite, thought to be of Silurian age occurs nearby. A metamorphic basement has been inferred at a depth of approximately 7 km beneath the orebody on the basis of geophysical data²⁰.

There is abundant evidence that the mineralization is sedimentary with early diagenetic additions and modifications^{29,30}. The stratigraphically lowest lens, number 5, was deposited first²⁸ and the main phase of mineralization terminated with deposition of

the stratigraphically highest number 1 lens. Some mineralization continued after the development of a submarine erosion surface with localized deposition of the pyritic Conglomerate Group Ore. The fact that mineralization extends into these Viséan sediments demonstrates the long-lived nature of this system ($>10^6$ years). Galena samples were taken from all five lenses to provide a comprehensive study of Pb isotopic variation within this deposit. In addition, stratigraphically less well-defined samples from the associated exhalative ore at Tynagh and Silvermines were analysed to provide a comparison with both the Navan data and previously published data for Irish mineral deposits^{31–33}. All data and details of analytical techniques are presented in Table 1.

In Fig. 2 the $^{206}\text{Pb}/^{204}\text{Pb}$ ratios of the galena samples are shown as a function of stratigraphic lens position (hence time). The most radiogenic signature ($^{206}\text{Pb}/^{204}\text{Pb} = 18.2$) persists as a regular feature throughout the orebody (Fig. 2). Cross-cutting veins (not plotted) are in the range of the more radiogenic lead isotopic compositions observed. Although isotopic variation is apparent on a 'lens scale' the replicate analyses in Table 1 (from separate portions of the same hand specimen) are remarkably uniform.

The only model which specifically predicts a change with time to less radiogenic Pb isotopic compositions is that of downward

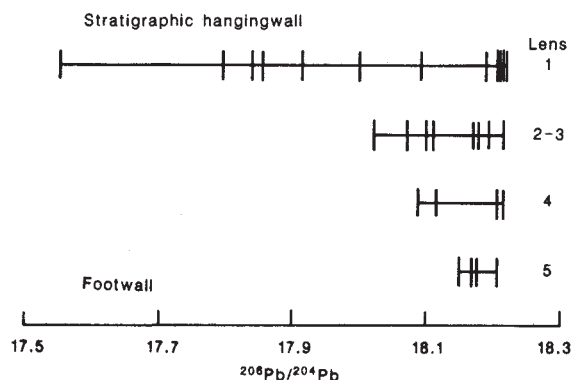


Fig. 2 Variation in $^{206}\text{Pb}/^{204}\text{Pb}$ in galenas with stratigraphic height in the Navan orebody.

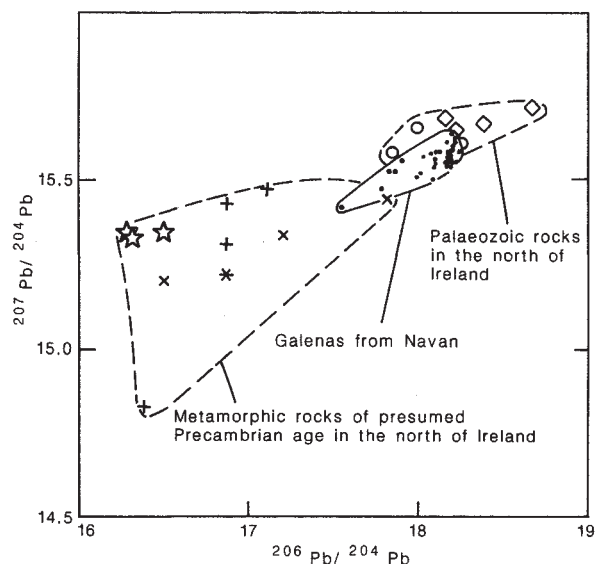


Fig. 3 Lead isotope data for Navan galenas compared with age corrected data (H. Mills, unpublished) for rock types in the north of Ireland. It is not thought that all these rock types underlie Navan but the data nonetheless provide strong support for the model shown in Fig. 4. Newry data on feldspars. The remainder are on whole rocks. ☆, Lewisian Inishtrahull; + late Proterozoic—Ox Mountains; ×, late Proterozoic—Tyrone; *, granulite facies xenolith; ◇, Lower Palaeozoic sediments—Navan; ○, Newry granodiorite; ●, Navan galenas.

excavating convection cells. The fact that this prediction is fulfilled does not on its own rule out the relevance of other available models. However we doubt that they are significant for the following (and other¹⁶) reasons. There are striking differences between the kinds of results that have been observed in the lead-zinc deposits of the Mississippi Valley, that is, highly variable and radiogenic Pb, and those observed at Navan. Furthermore, Mississippi Valley type deposits *sensu stricto* commonly exhibit very large interdeposit differences in isotopic composition which does not fit with the relatively uniform, systematic lead isotopic compositions found across Ireland as a whole³¹⁻³³. A contemporary igneous source for the lead does not seem very reasonable since it does not explain the temporal variations in lead isotopic composition; but more importantly no Carboniferous igneous rocks³⁴ of significance are known in the vicinity. The fluid inclusion temperatures in Irish SEDEX deposits (estimated maximum of 250 °C) are too high to be easily reconciled with seismic pumping or basin brine expulsion models¹⁶. Finally, the fact that the Navan orebody was formed

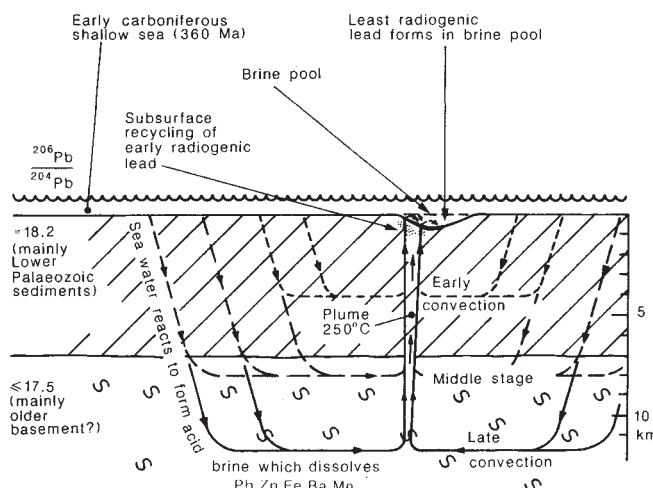


Fig. 4 Diagram illustrating the progress of a convection cell as it excavates downwards to tap older continental crust with less radiogenic lead which is transported upwards to form a giant exhalative orebody such as Navan.

early in the development of a sedimentary basin, is inconsistent with basin brine expulsion models.

The assumption that lead should be less radiogenic with depth in the crust is based on a generalized model that needs vindication by analysing local rock types. In Fig. 3 the lead isotope data for Navan mineralization is compared with unpublished data for a variety of potential type reservoirs in the northern half of Ireland. It can be seen that local Lower Palaeozoic sediments and 400 Myr old granites (as typified by the nearby Newry granodiorite) have lead isotopic compositions clustering around the radiogenic end of the spectrum of Navan results. In contrast, a variety of presumed Precambrian rocks in the northern half of Ireland have less radiogenic lead such that the Navan galenas span the two fields.

Based on this and geophysical²⁰ evidence a reasonable generalized 'lead isotopic stratigraphy' of the crust is suggested in Fig. 4 in which the Navan orebody is shown immediately underlain by Lower Palaeozoic sediments which in turn are underlain by a Precambrian basement at ~7 km depth. An explanation for the formation of the orebody in terms of deepening, expanding convection would be as follows. Carboniferous seawater penetrated the crust along fractures during periods of extension and the fluids spread out laterally at the point where penetration along fractures was no longer possible, that is, at the brittle/ductile transition. As the fluids circulated they were converted to acid brines which scavenged and mixed lead and other constituents from the uppermost crust. These buoyant hydrothermal fluids precipitated their constituents on the sea floor and within the sediments. The convection cell system advanced to greater depth with the brittle/ductile transition zone as the crust cooled.

The downward excavating convection model, on its own, does not explain the repeated tapping of lead with a radiogenic lead isotopic composition which persists as a signature throughout the evolution of the hydrothermal system. We consider that this lead is derived by the recycling in the uppermost crust of the lead homogenized and deposited during the formation of the lowest (and in fact the biggest) lens (lens number 5). Although we believe these processes are not unique to Navan, the time-resolved high-precision Pb isotope data are not yet available for other giant orebodies to determine whether this model of expanding convection cells is widely applicable.

We are indebted to Tara Mines for access to the Navan mine and drillcore, and to the chief geologist, Dermot Downing for his help. We would also like to thank our colleagues at Strathclyde University, SURRC and University of Michigan for

helpful comments and Alan Dickin and John Hutchinson for technical assistance. This work was carried out while two of us, H.M. and I.K.A. were in receipt of NERC grants. The isotope geology unit at SURRC is supported by NERC and the Scottish Universities.

Received 16 October 1986; accepted 26 February 1987.

- Halliday, A. N. & Mitchell, J. G. *Trans. R. Soc. Edin.* **74**, 1-14 (1983).
- Evans, A. M. in *Handbook of Stratatound and Stratiform Ore Deposits* Vol. 5 (ed. Wolf, K. H.) 231-255 (Elsevier, Amsterdam, 1976).
- Gustafson, L. B. & Williams, N. *Econ. Geol.* **876**, 139-178 (1981).
- Badham, J. P. N. *Trans. Instn Min. Metall.* **B90**, 70-76 (1981).
- McIntyre, D. G. in *Short Course in Sediment Hosted Lead-Zinc Deposits* (ed. Sangster, D. E.) 141-174 (Mineral. Assoc. Canada, Victoria, 1983).
- Lydon, J. W. in *Short Course in Sediment Hosted Lead-Zinc Deposits* (ed. Sangster, D. E.) 141-174 (Mineral. Assoc. Canada, Victoria, 1983).
- Morrissey, C. J., Davis, G. R. & Steed, G. M. *Trans. Instn. Min. Metall.* **B80**, 174-185 (1971).
- Evans, A. M. & Maroof, S. I. *Mining Mag.* **134**, 401-411 (1976).
- Campbell, F. A. & Ethier, V. G. *Miner. Deposita* **18**, 39-55 (1983).
- Finlow-Bates, T. *Geol. Jb.* **40**, 131-168 (1980).
- Russell, M. J. in *Implications of Continental Drift to the Earth Sciences* Vol. 1 (eds Tarling, D. H. & Runcorn, S. K.) 581-597 (Academic, London, 1973).
- Phillips, W. J. *Trans. Instn Min. Metall.* **B92**, 102 (1983).
- Gulson, B. L., Perkins, W. G. & Mizan, K. J. *Econ. Geol.* **78**, 1466-1504 (1983).
- Sibson, R. H., Moore, J. McM. & Rankin, A. H. *J. geol. Soc. Lond.* **131**, 653-659 (1975).
- Russell, M. J. *Trans. Instn Min. Metall.* **B87**, 168-171 (1978).
- Russell, M. J. in *Short Course in Sediment Hosted Lead-Zinc Deposits* (ed. Sangster, D. E.) 251-282 (Mineral. Assoc. Canada, Victoria, 1983).
- Due, B. R. & Delevaux, M. H. *Econ. Geol.* **67**, 409-425 (1972).
- Hart, S. R., Shimizu, N. & Sverjensky, D. A. *Econ. Geol.* **76**, 1873-1878 (1981).
- Doe, B. R. in *Lead Isotopes* 1-137 (Springer, New York, 1970).
- Brown, C. & Williams, B. J. *geol. Soc. Lond.* **142**, 1059-1075 (1985).
- Derry, D. R., Clark, G. R. & Gillatt, N. *Econ. Geol.* **60**, 1218-1237 (1965).
- Russell, J. M. *Trans. Instn Mon. Metall.* **B84**, 128-133 (1975).
- Boast, A. M., Coleman, M. L. & Halls, C. *Econ. Geol.* **76**, 27-55 (1981).
- Taylor, S. & Andrew, C. J. *Trans. Instn Min. Metall.* **B87**, 111-124 (1978).
- Taylor, S. *Econ. Geol.* **79**, 529-548 (1984).
- Banks, D. A. *Nature* **313**, 128-131 (1985).
- Boyce, A. J., Coleman, M. L. & Russell, M. J. *Nature* **306**, 545-550 (1983).
- Andrew, C. J. & Ashton, J. H. in *Mineral Exploration in Ireland, Progress and Development 1971-1981* (Irish Ass. for Economic Geologists, Dublin, 1982).
- Andrew, C. J. & Ashton, J. H. *Trans. Instn Min. Metall.* **B94**, 66-93 (1985).
- Anderson, I. K. *Abstr. Vol. Min. Dep. Study Group* (University of Strathclyde, 1985).
- Boast, A. M., Swainbank, I. G., Coleman, M. L. & Halls, C. *Trans. Instn. Min. Metall.* **B90**, 115-119 (1981).
- Caulfield, J. B., Lethusay, A. P. & Rye, D. M. in *Geology and Genesis of Mineral Deposits in Ireland* (eds Andrew, C. J. et al.) (Irish Ass. for Economic Geology, Dublin, 1986).
- O'Keeffe, W. G. *Geology and Genesis of Mineral Deposits in Ireland* (eds Andrew, C. J. et al.) 617-624 (Irish Ass. for Economic Geology, Dublin, 1986).
- Smedley, P. L. *Earth planet. Sci. Lett.* **77**, 113-128 (1986).
- Arden, J. W. & Gale, N. H. *Analyt. Chem.* **46**, 2-9 (1974).

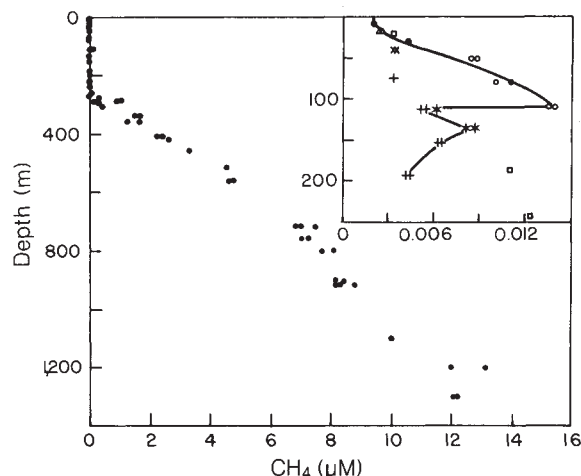


Fig. 1 Methane distribution in western Cariaco Basin during February and March 1986. Inset: expanded scale for surface layer. Different symbols represent different casts. Methane concentrations determined using vacuum extraction of dissolved gas from water sample²³ followed by injection of 100 to 500 μ l gas into Carle AGC 211 GC equipped with a flame ionization detector and 1/8-inch outer diameter, 5-foot stainless-steel column containing 60/80 mesh molecular sieve 5A. Precision and accuracy of concentration data are 3-5%. Equilibrium surface methane concentration $C_{eq} = \beta \chi_i^A (P - p^*)$ where C_{eq} is the equilibrium gas concentration, β , the Bunsen solubility coefficient²⁴, χ_i^A the partial pressure of gas in air ($=1.7$ p.p.m.v.)^{8,9}, P , the barometric pressure and p^* , the water vapour pressure. Air-sea flux $= K_L (C_{meas} - C_{eq})$ (ref. 9) where $K_L = 3.8$ m per day is the liquid phase gas transfer coefficient at a wind speed of 8.5 m s⁻¹ and C_{meas} = *in situ* gas concentration at sea surface. Diffusional fluxes at concentration gradients $= K_v (\Delta C / \Delta z)$ where K_v is the vertical eddy diffusivity calculated from conductivity, temperature and depth (CTD) data as a function of the density gradient¹¹. $K_v = 0.20$ cm² s⁻¹ in the 50-80 m interval and 0.6 (from model⁷) or 0.78 (from density gradient) cm² s⁻¹ in the oxic/anoxic interface region (260-290 m). $\Delta C / \Delta z$ = the methane concentration gradient.

Methane oxidation and methane fluxes in the ocean surface layer and deep anoxic waters

B. B. Ward*, K. A. Kilpatrick*, P. C. Novelli† & M. I. Scranton†

* Institute of Marine Resources, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA

† Marine Science Research Center, State University of New York, Stony Brook, Long Island, New York 11794, USA

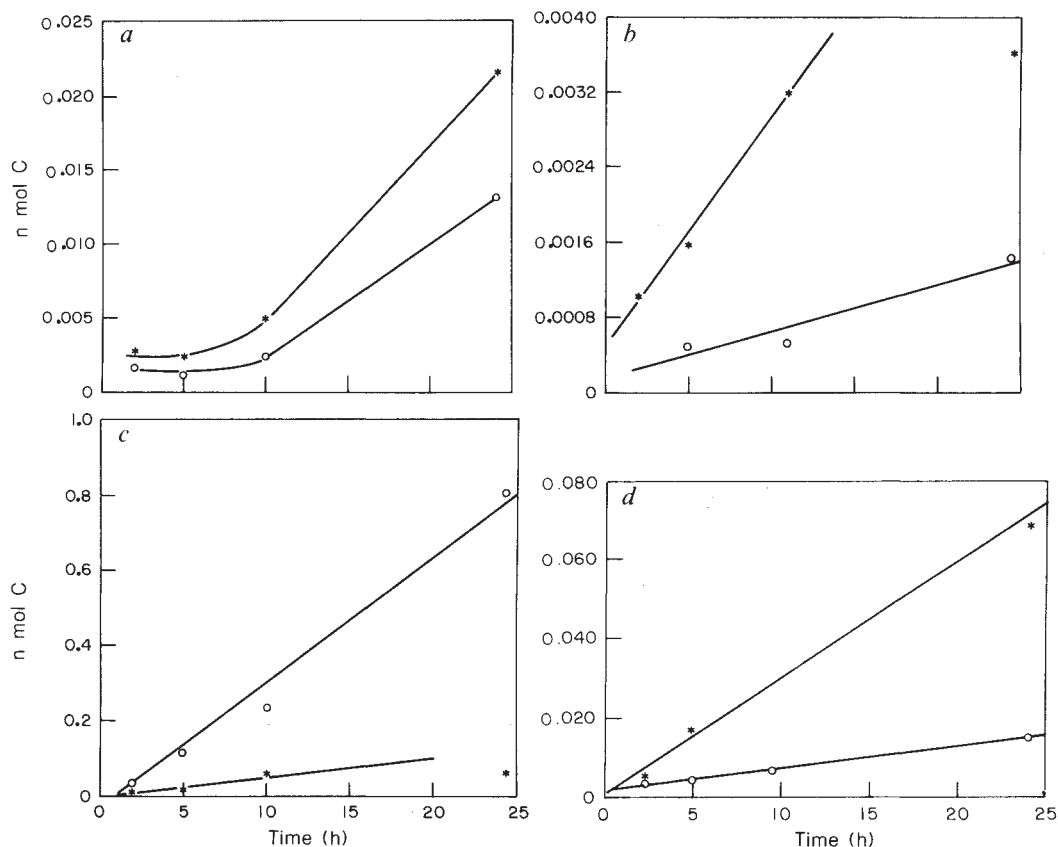
Methane is supersaturated in sea water, and is typically at its maximum concentration in near-surface waters, which could support a significant sea-air flux. The magnitude and variability of the flux depends on the mechanisms which produce and consume methane in sea water. Here, we compare measured biological oxidation rates of methane with the diffusional fluxes computed from concentration gradients in the surface layer of the ocean, and show that oxidation of methane in sea water is a mechanism which modulates the flux of methane from marine waters to the atmosphere. Methane fluxes and oxidation rates were investigated in surface waters, at the oxic/anoxic interface and in deep anoxic waters of the Cariaco Basin. Measured oxidation rates were

equivalent to 5% of the methane flux into oxygenated waters from the methane-rich deep waters and 10% of the flux into surface waters from the subsurface methane maximum. Thus oxidation was not sufficient to prevent a net sea-air flux. The total methane oxidation rate in the basin amounted to 1.5% of total primary production in the surface layer. Only a small fraction of oceanic primary production would be required to cycle through the methane pool to support the global atmospheric flux from the ocean¹.

A slight supersaturation of methane persists in sea water, even in well-oxygenated surface waters in the absence of known production mechanisms. Subsurface maxima are commonly observed²⁻⁵ but net fluxes between the ocean and the atmosphere of only several tens of nmol cm⁻² yr⁻¹ have been estimated. The diffusive flux away from the subsurface maximum is generally smaller than the flux at the air/sea interface, implying an *in situ* production term in the surface layer⁴. The shape of the subsurface maximum implies that the flux of methane to the atmosphere is controlled by *in situ* biological consumption processes.

In anoxic basins, there is an additional source of methane from anaerobic metabolism in the sediments. This is the case in the Cariaco Basin⁶; its oligotrophic surface waters cap a deep anoxic basin, separated from the Caribbean Sea by a sill at 150 m. The oxic/anoxic interface is generally found between 260 and 300 m, although the precise depth has varied over time^{6,7}. Below the interface, methane and sulphide accumulate to very high concentrations. Anoxic basins and sediments constitute large methane pools, which could contribute to the sea-air flux.

Fig. 2 Time course of methane oxidation at four depths. O, cell carbon; *, CO₂. *a*, Oxidation at 10 m; note initial lag. *b-d*, Linear regressions used to determine rates; 60 m ($r > 0.96$), 240 m ($r > 0.96$) and 700 m ($r > 0.99$). Samples collected in 5- or 30-litre Niskin bottles and transferred to incubation flasks (250 ml) without bubbles with overflowing. Four flasks per experiment were sealed without bubbles and the tracer methane (55 mmol mCi⁻¹) added by syringe. Time point procedure: add 0.1 ml 10 M NaOH, mix and filter through a 0.3 µm Millipore membrane filter. Transfer filtrate to 1.0 litre bottle and seal with phenethylamine soaked glass fibre filter suspended from lid. Inject 1.0 ml concentrated H₂SO₄; incubate 1 h; mix periodically. The glass fibre filter (containing labelled CO₂) and the membrane filter (containing labelled particulate material) from each time point were dissolved in a beta-phase scintillation cocktail and counted in an LS-100 Beckman liquid scintillation counter upon return to the laboratory (3 to 4 weeks). Precipitate formed upon the addition of base caused 0.33% or less of the dissolved ¹⁴CO₂ to be caught on the particulate filter. CO₂ capture efficiency in the acidification step was 50%. Data have been corrected for these controls. Substrate concentration was not significantly altered during incubation; measurement of methane (by gas chromatography following the cruise) in unfiltered subsamples detected no change in total methane among time points. Label incorporation was calculated as $[\text{d.p.m.}_{\text{sample}} - \text{bkg} / \text{d.p.m.}_{\text{initial}}] \times [\text{initial CH}_4 \text{ concentration}]$ (bkg, background). Rates were calculated by linear regression. Changes in gas concentrations induced by addition of a 250 µl bubble of tracer gas were calculated by the equation of Rudd and Hamilton¹⁴.



Dissolved methane concentrations were determined by combining a vacuum extraction method with gas chromatography (Fig. 1). The main features of the methane distribution are low concentrations in the surface layer and rapidly increasing concentrations at the interface, with maximum concentrations of 12.1 µM at 1,300 m (Fig. 1). When seen with greater resolution (Fig. 1, inset) at least one maximum can be detected within the surface layer, separated from the high concentrations below the interface by a minimum around 200 m. Surface waters were very near equilibrium with the atmosphere; measured concentration at 10 m was 2.0 nM and the calculated surface equilibrium concentration^{8,9} was 1.93 nM. Thus a net sea-air flux of only about 0.023 nmol cm⁻² per day was calculated¹⁰, using a gas transfer coefficient of 3.77 m per day at a wind speed of 8.5 m s⁻¹.

The flux¹¹ into the surface layer from the 110-m subsurface maximum (0.020 nmol cm⁻² per day calculated from the concentration gradient and a vertical eddy diffusion coefficient of 0.20 cm² per day) implies that diffusion alone could supply the air-sea flux if no *in situ* consumption occurs. The flux into the oxygenated layer across the oxic/anoxic interface was estimated to be 5.02 to 6.53 nmol cm⁻² per day, depending on the diffusion coefficient (Fig. 1)¹¹. Differences between the present calculations and Reeburgh's¹² estimates of 0.055 to 0.71 nmol cm⁻² per day may be due to important differences in sample storage and analysis, and to temporal change in the deep water methane concentration⁷. This flux may vary seasonally as well; our measurements were made during a windy period. Since the calculated flux at the oxic/anoxic interface far exceeds the

estimated fluxes away from the subsurface maximum or across the air/sea interface, the existence of a large *in situ* sink for methane between 110 m and the oxic/anoxic interface is suggested. In the flux calculations, we have assumed simple vertical diffusive transport of gas. Evidence of spatial and temporal variability in the methane data is present in Fig. 1, and may represent lateral transport of methane-rich water from near the coast. If significant amounts of methane are supplied by advection, our estimates of *in situ* production (see below) may be too high. The minimum in methane concentration above the oxic/anoxic interface implies that the deep reservoir is not the only source of methane to the surface layer.

Using biogenically produced ¹⁴C-methane (ref. 13) as a tracer for methane oxidation, we measured the rate of production of ¹⁴C-labelled particulate material and ¹⁴CO₂ in a depth profile in the Cariaco Basin. Rates were determined from timecourse experiments, in which four incubation flasks were inoculated at time zero, and one flask was terminated at each time point between 2 and 24 h (Fig. 2). Incubations were performed without oxygen contamination at *in situ* temperatures in the dark. Rates were calculated from the linear regression of carbon versus time; only incubations which yielded a minimum of three points defining a statistically significant slope ($P < 0.05$) are included in the following calculations. Single endpoint incubations were used to assess the dependence of oxidation rate on substrate concentration. Linear dependence, with an intercept not different from zero, was observed (Fig. 3). Rates calculated from the time course experiments were corrected to *in situ* values by

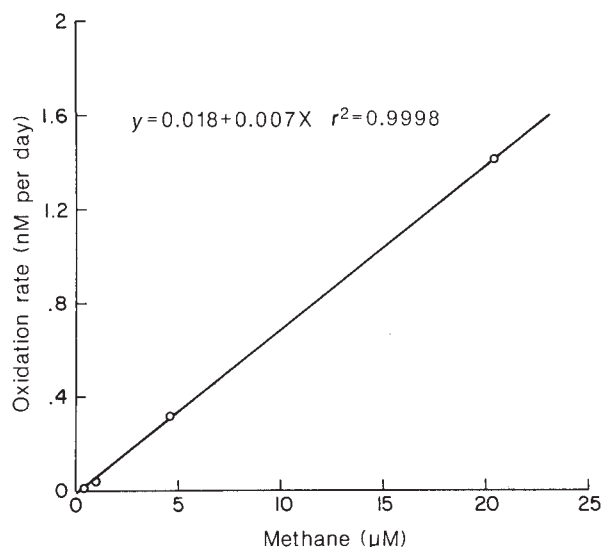


Fig. 3 Relationship between methane oxidation rate and methane concentration for a sample from 50 m.

assuming oxidation rates at all depths were first order with respect to substrate concentration^{14,15}.

Three patterns of methane oxidation were observed (Fig. 2): (1) in the oxygenated surface layer (Fig. 2a, b), cell production constituted 9.2% of total methane oxidation (± 2.3 , $N=6$; including only depths where both CO_2 and cell production yielded significant slopes). In the 10 m sample (and 25 m sample, not shown), a lag was apparent in the time courses, suggesting that methane oxidation may have been induced in these samples. If the microorganisms responsible for methane oxidation are light sensitive¹⁶, after some time, incubation in the dark may have removed light inhibition resulting in linear methane oxidation in the later part of the time course (Fig. 2a). Total oxidation rate was on the order of 10^{-4} nM per day with specific oxidation rate (nM CH_4 oxidized/nM CH_4 ambient) of 4.1×10^{-5} per day ($\pm 1.4 \times 10^{-5}$). (2) At the oxic/anoxic interface, total oxidation rates were 10^{-2} nM per day with specific rates up to 2.5×10^{-3} per day. Cell production averaged 61.7% (± 8.8 , $N=4$) of the total methane oxidation (Fig. 2c). Most aerobic oxidation probably takes place in a very thin layer, although the increase in rate approaching the interface may begin at a depth as shallow as 150 m (Fig. 4). (3) In the deep anoxic water, CO_2 production again exceeded cell production (Fig. 2d), which was 9.0% of the total (± 2.0 , $N=3$). Total oxidation rates were 0.4 nM per day and specific oxidation rates averaged 4.2×10^{-5} per day ($\pm 1.3 \times 10^{-5}$).

The depth distribution of methane oxidation rates in the Cariaco Basin (Fig. 4) resembles that found in dimictic lakes¹⁴, with a thin layer of high aerobic oxidation rates occurring at the oxic/anoxic interface. However, a major difference from freshwater systems was that the highest oxidation rates, 0.44 nM per day, were observed in the deep anoxic water. Except that the rate process in the deep water occurred in the absence of oxygen and in the presence of sulphide, it is similar in cell/ CO_2 ratio and efficiency (substrate specific oxidation rate) to the process in the surface layer, and distinct from that at the interface. Oxidation at the interface was much more efficient per unit substrate and produced a much greater proportion of cell material. The apparent efficiency of cell production from methane in the interface samples exceeds that reported for methanotrophs in pure culture (about 50%)¹⁷ and for methane oxidation in lakes (22% to 33%)¹⁸. The values reported here may be slightly higher than the net assimilation efficiency since they were calculated as a ratio of rates during linear methane oxidation over relatively short timescales. Some of the carbon

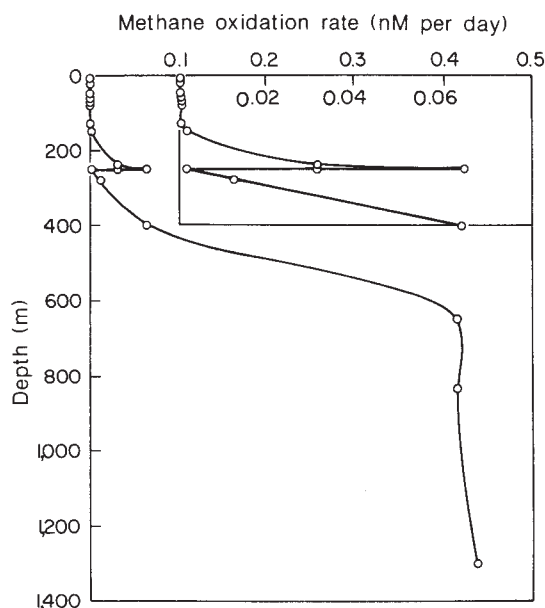


Fig. 4 Depth distribution of total methane oxidation rate (cell + CO_2 production). Rates at ambient methane concentration were calculated from measured rates assuming a first-order reaction.

immediately fixed into cell material may be respired on slightly longer timescales. These high ratios may also imply that microorganisms other than conventional methanotrophs, which occur in lakes, are responsible for some methane oxidation in sea water¹⁹.

Integrated oxidation rates ($0.002 \text{ nmol cm}^{-2}$ per day) between the subsurface methane concentration maximum (110 m) and the sea surface were 10% of the calculated flux away from the maximum. Total loss of methane from the surface layer (0–150 m; oxidation plus the air-sea flux = $0.006 + 0.023 \text{ nmol cm}^{-2}$ per day) is balanced by production, which must be at least $0.009 \text{ nmol cm}^{-2}$ per day (total loss minus diffusive supply = $0.029 - 0.02 \text{ nmol cm}^{-2}$ per day). This is a rough estimate, since the measurement intervals for the oxidation rates and diffusional fluxes did not coincide exactly. While this estimate of methane production rate is derived indirectly, it allows us to assess the relative importance of *in situ* production, which is very difficult to measure directly because the immediate precursors and biochemical reactions involved are presently unknown.

Integrated methane oxidation in the interval 150–280 m ($0.279 \text{ nmol cm}^{-2}$ per day) was about 5% of the estimated diffusional input indicating that the anoxic reservoir can, at least at times, contribute to the surface layer methane pool, and perhaps to the atmospheric flux. Our rate measurements may underestimate the intensity of the oxidation rate at the interface, given the small vertical extent of this layer. In addition, model calculations⁷ suggest that high oxidation rates may occur in this depth interval as mixing occurs across the interface in response to seasonal or episodic physical forcing. At the time of our study, the Cariaco Basin was a small net source of methane to the atmosphere because *in situ* oxidation rates were not rapid enough to consume the diffusional flux across either the oxic/anoxic or the sea/air interface. However, the minimum in methane concentration between the subsurface maximum and the interface is evidence that most methane produced in the anoxic region is consumed near the interface, rather than diffusing unimpeded into the surface layer. Our measured rates must be underestimates of the long-term mean consumption of methane at the interface due to temporal and spatial variability in the *in situ* methane oxidation rate.

Integrated methane production estimated for the surface layer ($0.009 \text{ nmol cm}^{-2}$ per day) is small compared with primary

production integrated over the 100 m photic zone (2.4×10^3 nmol cm^{-2} per day measured on the same cruise using *in situ* incubations with the ^{14}C -bicarbonate method; results will be reported in detail elsewhere). Production of organic matter by incorporation of methane between 240 and 250 m (the aerobic oxidation maximum) was 0.026 nmol C cm^{-2} per day, equivalent to 0.001% of integrated primary production. Integrated methane oxidation between 280 and 1,300 m was 33.8 nmol C cm^{-2} per day, equivalent to 1.4% of primary production. Total methane oxidation, integrated from surface to bottom waters, consumed 1.5% of daily primary production. Methane oxidation and primary production probably vary seasonally²⁰, and while we are not implying an immediate dependency of methane oxidation rates at 1,000 m on surface layer photosynthesis, the comparison demonstrates the relative importance of each process. The data imply that very little of the carbon fixed in the oligotrophic-surface ocean cycles through the methane pool. If our data are representative of other oligotrophic areas, they suggest that the total ocean-atmospheric flux of methane can be supported by less than 0.1% of primary production.

Of particular interest in the marine environment are the observations of methane oxidation in both the surface oxygenated layer and the deep anoxic layer. The rate data from the time course experiments, and the response of oxidation rate to increased substrate concentration are evidence for the presence of microorganisms capable of methane oxidation.

A minimum in the oxidation rate occurred at the oxic/anoxic interface. This inflection in the depth distribution of oxidation rates, and the striking differences in substrate specific rate and cell/ CO_2 ratio on either side of the minimum is consistent with the presence of different types of microorganisms in different parts of the water column. Below this minimum, significant methane oxidation rates were found in the anoxic deep water, with rates increasing between 280 m and 1,300 m. Based on a steady-state model, Reeburgh¹⁹ predicted that anoxic methane oxidation should occur in the deep waters at rates of 0.003 to 0.042 nmol l^{-1} per day. Our measured rates are 0.06 nmol l^{-1} per day at 400 m to 0.44 nmol l^{-1} per day at 1,300 m. Recent studies indicate that the steady-state assumption is invalid for the Cariaco Trench⁷, so the apparent agreement between his upper estimate and our lower one may be fortuitous. While the occurrence of anaerobic methane oxidation is a highly controversial topic, several facts support our conclusion that anaerobic oxidation was observed in our samples. (1) The functioning of contaminating aerobic organisms in anaerobic conditions is not consistent with the observed minimum in total rates just below the interface. (2) Zero time samples were less than 10% of the 2 h samples and linearity in both $^{14}\text{CO}_2$ and ^{14}C -cell production rates were observed. (3) High sulphide concentrations in the deep samples are incompatible with the physiology of conventional aerobic methanotrophs. (4) Oxygen contamination of the samples during handling was minimal; good micro-oxygen and sulphide assays from the sampling bottles provide evidence that they remained sealed even on deck until opened for subsampling. Minor contamination of samples with oxygen during manipulations cannot be ruled out in every case, and while abiological consumption of oxygen by sulphide is slow at low oxygen concentrations²¹, the reasoning above leads us to conclude that the observed oxidation process was not an oxygen dependent reaction.

Using an inhibitor approach to estimate potential methane oxidation in incubations up to ten days, Lidstrom²² detected a qualitatively similar pattern of methane oxidation in a partially anoxic fjord. Methane oxidation was detected only at the oxic/anoxic interface, where the process was aerobic, and in the deep water below the interface, where the oxidation proceeded in the presence of sulphide at rates up to $10 \mu\text{M}$ per day. The tracer method used here is more sensitive and permitted detection of significant rates in oxygenated surface waters as well.

This research was supported by the US NSF and NASA. Methanogens for biogenic methane production were kindly supplied by Dr L. Daniels. We thank the crew of the RV *Iselin* and O. C. Zafiriou, chief scientist.

Received 5 September 1986; accepted 27 February 1987.

- Sieler, W. in *Current Perspectives in Microbial Ecology* (eds Klug, M. J. & Reddy, C. A.) 468-477 (American Society for Microbiology, Washington DC, 1984).
- Swinnerton, J. W., Linnenbom, V. J. & Cheek, C. H. *Envir. Sci. Technol.* **3**, 836-838 (1969).
- Lamontagne, R. A., Swinnerton, J. W., Linnenbom, V. J. & Smith, W. D. *J. geophys. Res.* **78**, 5317-5324 (1973).
- Scranton, M. I. & Brewer, P. G. *Deep Sea Res.* **24**, 127-138 (1977).
- Burke, R. A., Reid, D. F., Brooks, J. M. & Lavoie, D. M. *Limnol. Oceanogr.* **28**, 19-32 (1983).
- Richards, F. A. *Rev. Oceanogr. mar. Biol.* **13**, 11-67 (1975).
- Scranton, M. I., Sayles, F. L., Bacon, M. P. & Brewer, P. G. *Deep Sea Res.* (in the press).
- Khalil, M. A. K. & Rasmussen, R. A. *Science* **232**, 56-58 (1986).
- Rinsland, C. P., Levine, J. S. & Miles, T. *Nature* **318**, 245-249 (1985).
- Hartman, B. & Hammond, D. E. *Hydrobiologica* **129**, 59-68 (1985).
- Sarmiento, J. L., Feely, H. W., Moore, W. S., Bainbridge, A. E. & Broecker, W. S. *Earth planet. Sci. Lett.* **32**, 357-370 (1976).
- Reeburgh, W. S. *Earth planet. Sci. Lett.* **28**, 337-344 (1976).
- Daniels, L. & Zeikus, J. G. *J. Label. Comp. Radiopharm.* **20**, 17-24 (1983).
- Rudd, J. W. M. & Hamilton, R. D. *Arch. Hydrobiol.* **75**, 522-538 (1975).
- Harris, S. M. & Hanson, R. E. *Limnol. Oceanogr.* **25**, 412-421 (1980).
- Sieburth, J. McN., Johnson, P. M., Eberhardt, M. A. & Sieracki, M. E. *Eos* **64**, 1054 (1983).
- Anthony, C. *The Biochemistry of Methylophony* (Academic, London, 1982).
- Hanson, R. S. *Adv. appl. Microbiol.* **26**, 3-39 (1980).
- Jones, R. D., Morita, R. Y. & Griffiths, R. P. *Mar. Ecol. Prog. Ser.* **17**, 259-269 (1984).
- Castellvi, J. & Alvarez, E. F. *Biol. Inst. Oceanogr. Univ. Oriente* **17**, 9-29 (1978).
- Lidstrom, M. E. *Limnol. Oceanogr.* **28**, 1247-1251 (1983).
- Millero, F. J., Hubinger, S., Fernandez, M. & Garnett, S. *Envir. Sci. Technol.* (in the press).
- Herr, F. L. & Barger, W. R. *J. geophys. Res.* **83**, 6199-6205 (1978).
- Yamamoto, S., Alcauskas, J. B. & Crozier, T. E. *J. chem. Engng Data* **21**, 78-80 (1976).

Abolition of the expression of inhibitory guanine nucleotide regulatory protein G_i activity in diabetes

Debra Gawler, Graeme Milligan, Allen M. Spiegel*, Cecilia G. Unson† & Miles D. Houslay‡

Molecular Pharmacology Group, Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, UK

* Metabolic Diseases Branch, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

† Department of Biochemistry, The Rockefeller University, New York, New York 10021, USA

Many cell-surface receptors for hormones appear to exert their effects on target cells by interacting with specific guanine nucleotide binding regulatory proteins (G -proteins) which couple receptors to their second-messenger signal generation systems. A common intracellular second messenger, which is used by many hormones, is cyclic AMP. This is produced by adenylate cyclase, whose activity is controlled by two G -proteins, G_s which mediates stimulatory effects and G_i inhibitory effects on adenylate cyclase activity¹. In liver, the hormone glucagon increases intracellular cAMP concentrations by activating adenylate cyclase by a G_s -mediated process. This effect of glucagon is antagonised by the hormone insulin, although the molecular mechanism by which insulin elicits its actions is obscure. However, insulin receptors exhibit a tyrosyl kinase activity² and appear to interact with G -proteins^{2,3}, perhaps by causing phosphorylation of them⁴. In type I diabetes, circulating insulin levels are abnormally low, giving rise to gross perturbations of metabolism as well as to a variety of complications such as ionic disturbances, neuropathies of the nervous system, respiratory and cardiovascular aberrations and predisposition to infection⁵. We show here that experimentally-induced type I diabetes leads to the loss of expression of G_i in rat liver. As it has been suggested that G_i may couple receptors to K^+ -channels^{6,7} as well as mediating the inhibition of adenylate cyclase, aberrations in the control of expression of this key regulatory protein in type I diabetes may be expected to lead to pleiotropic effects.

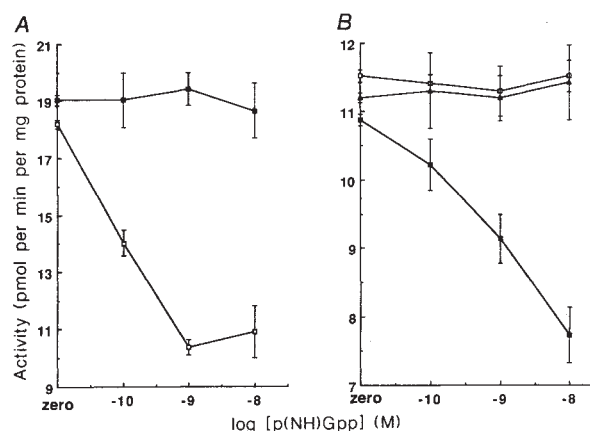


Fig. 1 Detection of functional G_i by inhibition of forskolin-stimulated adenylate cyclase activity using guanylyl 5'-imidodiphosphate. The inhibition of forskolin (10^{-4} M)-stimulated adenylate cyclase activity by low concentrations of p(NH)ppG is shown for hepatocyte membrane preparations from control animals (A) which either had (■) or had not (□) been pre-treated with pertussis toxin (100 ng ml^{-1} for 1 h), and for diabetic animals (B), as streptozotocin-diabetic (▲), alloxan-diabetic (□) and streptozotocin-diabetic animals undergoing insulin therapy (■). Values are means, (bars) for five different animals ($n = 5$) using triplicate adenylate cyclase assays. The basal adenylate cyclase activities of the various states are given in Table 1.

Methods. Male Sprague-Dawley rats (200–270g) were used. Diabetes was induced¹⁸ using one intraperitoneal injection of either streptozotocin (50 mg kg^{-1} ; 0.3 ml per animal) in sterile citrate buffer pH 4.0, or alloxan (130 mg kg^{-1} ; 0.3 ml per animal) in sterile 0.9% saline solution. Urine glucose was monitored (Diabur-Test 5000 kit, Boehringer) and animals were killed when they were clearly diabetic (7 days later with streptozotocin, 3 days later with alloxan). Animals were only confirmed as diabetic if glucose was detected in the urine and if the blood glucose concentration was $>12 \text{ mM}$ at the time of killing. Normal animals had a blood glucose concentration of $5.3 \pm 0.4 \text{ mM}$ and diabetic animals one of $17.0 \pm 1.4 \text{ mM}$. In some instances, diabetic animals (glucosuria), 72 h after streptozotocin treatment, were treated with insulin-zinc suspension (10 i.u. d^{-1}) injected subcutaneously each evening for 7 days, by which time glucosuria was not evident in the insulin-treated animals but was evident in those which had only been treated with streptozotocin. The blood glucose concentration of the insulin-treated animals was $3.8 \pm 0.2 \text{ mM}$. (Errors are s.d. for $n = 5$ separate animals).

Hepatocyte preparation¹⁵, incubation¹⁵ and membrane isolation¹⁶ was done as before. Similar yields of adenylate cyclase activity (range 40–50%) were obtained for all animals. Incubations of hepatocytes with pertussis toxin (100 ng ml^{-1}) were performed for 1 h at 37°C with frequent gassing (O_2/CO_2 , 95:5) and in the presence of 10 mM glucose to ensure cell viability during treatment^{8,17}. This treatment completely inactivated G_i (Fig. 1, refs 8 and 17) and effected ADP-ribosylation of G_i α -subunit *in situ*³.

Adenylate cyclase activity was assayed as described^{8,17}, using a binding protein for the determination of cAMP. Membranes were added at a concentration of $200\text{--}600 \mu\text{g ml}^{-1}$ ($20\text{--}60 \mu\text{g}$ per assay) and assays done at 30°C for 10 min, with initial rates of activity being taken for analysis. All errors are given as $\pm$ s.d. for $n = 5$ animals, with adenylate cyclase assays done in triplicate.

We⁸ and others⁹ have demonstrated that rat liver contains a functional G_i . Its activity can be assessed by demonstrating (Fig. 1A) the ability of low concentrations of the non-hydrolysable GTP analogue, p(NH)ppG (guanylyl 5'-imidodiphosphate) to inhibit adenylate cyclase activity which has been amplified by the diterpene, forskolin. In hepatocytes, low concentrations of p(NH)ppG inhibit forskolin-stimulated adenylate cyclase activity by some 35–45% (Fig. 1A). The functioning of G_i can be obliterated in various cells^{10–12}, including hepatocytes (Fig. 1A), by treatment of intact cells with pertussis toxin¹, which

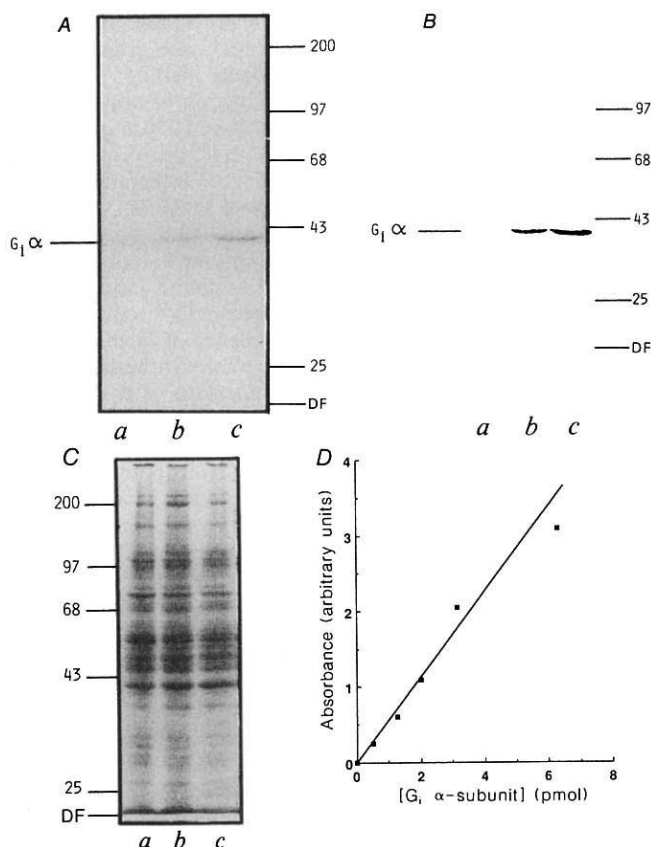


Fig. 2 Detection of G_i using the affinity purified polyclonal antiserum AS7. A, photograph and B, digitized image of an immunoblot showing tracks identifying G_i using the antibody AS7. Membranes are from nature animals in control (lane c), streptozotocin-diabetic (lane b) states: SF, dye front. Similar results were obtained using the antibody AS6. Each track was loaded with $300 \mu\text{g}$ of plasma membrane protein. Digitized scanning of the immunoblot on an Abatun 300 scanner was driven by an Apple Macintosh microcomputer (C-scan, v-1.0 software). C, Coomassie-blue stained tracks for the samples blotted in A and B. Size markers in a–c, M_r in thousands. These data are one typical set from an experiment that was done three times and show that there were not gross differences in the composition of the various membrane preparations examined. In all instances, contamination with mitochondrial endoplasmic reticulum markers³ was $<1\%$ and recovery of adenylate cyclase activity was similar (40–50% range). Identical amounts of protein ($50 \mu\text{g}$) were put on each gel track. D, Immunoreactivity of the affinity-purified antiserum AS7 against the purified α -subunit of G_i .

Methods. Immunoblotting of Percoll-purified³ rat-liver hepatocyte plasma membranes was carried out as described¹⁹ using *O*-dianisidine as a substrate for the peroxidase-linked second antibody. Affinity purified antisera AS6 and AS7, were prepared in rabbits using a synthetic decapeptide (KENLKDCGLF), representing the carboxy-terminus of the α -subunit of rod transducin^{20,21}, conjugated with keyhole limpet hemocyanin²². These antisera react strongly with purified bovine transducin and G_i α -subunits but not at all with purified bovine G_0 α -subunits. In D, increasing concentrations of G_i purified from bovine brain²³ were subjected to SDS-polyacrylamide gel electrophoresis (PAGE), under reducing conditions, and transferred to nitrocellulose. Affinity purified AS7 (1:100) was employed as first antibody and detected as above. Quantification with a Biorad Gel Scanner was driven by an Olivetti M24 microcomputer (1-D software) yielding a linear plot of absorption against G_i α -subunit concentration over the range examined. Such a plot was used to quantify G_i α -subunit in membranes. Using a similar range of G_0 concentration, this antibody (AS7) failed (less than 0.05 arbitrary units) to detect G_0 α -subunit. It is thus specific for the α -subunits of transducin and G_i . In liver it detects a 40K species which presumably reflects the 40K α -subunit of G_i which we can detect^{8,17} by ADP-ribosylation using pertussis toxin^{3,8}. Transducin does not occur in liver.

Table 1 Adenylate cyclase activity in normal and diabetic animals

Treatment	Basal activity (pmol per min per mg protein)	Forskolin stimulated activity (pmol per min per mg protein)	Maximal glucagon- stimulated activity (pmol per min per mg protein)	Stimulation by glucagon (activity ratio of glucagon- stimulated basal activities)	K_a -glucagon (nM)
Control	1.30 ± 0.09	18.2 ± 0.2	25.0 ± 2.1	19.2	14 ± 3
Alloxan-diabetic	0.72 ± 0.09	11.6 ± 0.1	24.3 ± 1.3	33.7	11 ± 2
Streptozotocin-diabetic	0.64 ± 0.02	11.2 ± 0.1	21.1 ± 0.6	33.0	14 ± 3
Insulin-treated streptozotocin-diabetic	0.68 ± 0.04	10.9 ± 0.1	17.2 ± 1.2	25.2	10 ± 5
Pertussis toxin- treated control	1.36 ± 0.12	19.0 ± 0.2	40.7 ± 3.0	29.9	15 ± 4
Pertussis toxin- treated streptozotocin diabetic	0.68 ± 0.04	10.7 ± 0.3	21.1 ± 2.1	31.1	12 ± 3
Pertussis toxin- treated alloxan diabetic	0.74 ± 0.04	ND	25.5 ± 2.5	34.4	11 ± 3
Pertussis toxin- treated insulin-treated streptozotocin diabetic	0.69 ± 0.05	ND	24.8 ± 2.5	35.9	12 ± 2

Hepatocyte membranes were prepared from rats as before¹⁵⁻¹⁷. Dose response curves were formulated using a range¹¹ of glucagon concentrations from 10^{-11} M to 10^{-7} M. From these a K_a value was calculated as the concentration giving half maximal activation by glucagon. The maximal glucagon-stimulated activity was that achieved using a saturating concentration (10^{-7} M). The ability of glucagon to stimulate basal adenylate cyclase activity was calculated by assessing the degree of stimulation it caused, that is the activity exhibited at 10^{-7} M glucagon as a proportion of the basal activity in the absence of glucagon. The diterpene forskolin magnifies basal adenylate cyclase activity, an effect shown with 10^{-4} M forskolin. Data for pertussis toxin-treated cells are from membranes prepared from cells which had been pretreated with pertussis toxin (100 ng ml⁻¹) for 1 h at 37 °C. Data is given with s.d. for five experiments, ($n = 5$), using different animals with triplicate adenylate cyclase assays.

causes the ADP-ribosylation of the α -subunit of G_i , which has relative molecular mass 40,000 (M_r 40K). Indeed, in hepatocytes, it has been shown that pertussis toxin causes the ribosylation of a single 40K species, representing the α -subunit of G_i , with concomitant loss of G_i function^{8,9}. However, when we used hepatocyte membranes from diabetic rats, low concentrations of p(NH)ppG failed to inhibit forskolin-stimulated adenylate cyclase activity (Fig. 1B). This could indicate the absence of functional G_i in liver membranes from diabetic rats. Such an effect was independent of whether alloxan or streptozotocin was used to induce diabetes (Fig. 1B). We noted that this effect could be reversed by treating streptozotocin-diabetic animals with insulin to reverse the hyperglycaemic state (Fig. 1B).

Diabetes also seemed to reduce the basal activity of adenylate cyclase when measured directly or when amplified using the diterpene, forskolin (Table 1). Indeed, insulin-treatment of streptozotocin-diabetic animals did not seem to reverse this effect which, together with the fact that pertussis toxin treatment did not alter basal adenylate cyclase activity, implies that the change in basal activity observed in the diabetic states induced by both alloxan and streptozotocin was not due directly to any change in G_i (Table 1). It is likely that the rather large reduction in the basal activity of adenylate cyclase caused by the diabetic state could result from altered expression of the catalytic unit of adenylate cyclase. Although alterations in total protein content of the membranes and small changes in the relative purity of the membrane fractions could affect this value, the similar yields of activity obtained and similar electrophoretic profiles (Fig. 2) indicate that this is probably not a major contributing factor. Thus diabetes may affect the expression of both G_i and adenylate cyclase itself.

Treatment of various cells with pertussis toxin, to obliterate functional G_i , has been shown to enhance the effect of stimulatory hormones on adenylate cyclase activity, presumably by removing the inhibitory input^{8,11,12} due to G_i . This can be seen using rat hepatocytes where pre-treatment with pertussis toxin (Table 1) markedly augmented the ability of glucagon to stimulate adenylate cyclase activity. However, pertussis toxin treat-

ment of hepatocytes from diabetic animals failed to enhance the degree of stimulation of adenylate cyclase that could be achieved by glucagon (Table 1). The reduction in basal specific activity of adenylate cyclase in the diabetic state militates against using a direct comparison of specific activities of the glucagon-stimulated state for assessing the potency of this hormone to stimulate adenylate cyclase. Thus, to allow a direct comparison, the potency of stimulation is expressed as a ratio of the glucagon-stimulated to basal activities, that is, as a degree of stimulation (Table 1). Thus, using membranes from diabetic animals we observed that glucagon exerted a far greater stimulatory effect on adenylate cyclase activity than was seen using hepatocyte membranes from control animals (Table 1). Indeed, the degree of the enhanced stimulatory effect observed when using membranes from diabetic animals compared well with that using membranes prepared from the hepatocytes of control animals that had been pre-treated with pertussis toxin to obliterate G_i functioning (Table 1). These data indicate that identical effects, in enhancing the ability of glucagon to stimulate adenylate cyclase, could be elicited by the removal of functional G_i (Fig. 1) either by pertussis toxin treatment or by inducing diabetes using streptozotocin or alloxan. Consistent with the removal of functional G_i in the diabetic state is therefore the failure of pertussis toxin treatment to alter activation by glucagon of adenylate cyclase in membranes from diabetic animals. Such enhanced activation of adenylate cyclase, in the diabetic state, can be reversed by treatment of the animals with insulin (Table 1), an action which has previously been shown to restore functional G_i in these membranes (Fig. 1). We did not observe any gross change in the K_a values for activation by glucagon, which were all in the range 10–15 nM for membranes prepared from all the various animals and before and after treatment with pertussis toxin.

The carboxy-terminus of the α -subunit of rod transducin is highly homologous^{13,14} to that of G_i . The antibodies AS6 and AS7, which we have raised against a synthetic decapeptide representing this region, can recognise the α -subunits of both transducin (M_r 39K) and G_i (M_r 40K) specifically. They do not recognise G_o , a G-protein which, like G_i , is also a substrate for

ribosylation by pertussin toxin. However, transducin, unlike G_i , is found only in retinal rods and cones. Thus these antibodies provide specific tools for identifying G_i in liver and can be used to quantify the G_i in membrane preparations. Immunoblotting of hepatocyte plasma membranes identified the single species at 40K, representing the α -subunit of G_i (Fig. 2). Using pure bovine brain G_i as a standard, we were able to determine that the amount of G_i α -subunit in these membranes was 7.5 ± 0.3 pmol per mg plasma membrane protein. However, in hepatocyte membranes from streptozotocin-diabetic animals the amount of G_i detected fell dramatically to 0.5 ± 0.2 pmol per mg plasma membrane protein. Treatment of streptozotocin-diabetic rats with sufficient insulin to normalize their blood glucose concentrations (Fig. 1) increased the amount of G_i in hepatocyte membranes to 4.1 ± 0.4 pmol per mg plasma membrane protein (data given as $\pm$ s.e.m. for five different animals). Thus, the regime of insulin treatment used returned liver plasma membrane concentrations of G_i to above 50% of those expressed by the control, adult animals. This value is not inconsistent with the degree of reversal observed when measuring the functional parameters of either inhibition of adenylate cyclase activity by p(NH)ppG or the potentiation of activation of adenylate cyclase activity by glucagon (Fig. 1; Table 1). These data indicate that in diabetes there is an actual reduction in G_i protein in liver. We consider it unlikely that any modification, such as, phosphorylation of G_i , could have prevented the detection of G_i by immunoblotting. First, there are no potential sites for phosphorylation within the sequence of the G-protein which these antibodies recognise; second, similar results were obtained using two different polyclonal antibody preparations (AS6, AS7) and thirdly G_i ADP-ribosylated by pertussis toxin treatment can be detected as effectively as native G_i by these two antibody preparations (data not shown). This is despite the fact that the site of ADP-ribosylation is within the sequence of amino acids forming the peptide used to raise these antibodies.

We suggest that the expression of G_i in liver is regulated either directly by circulating insulin concentrations or indirectly as a consequence of changes in cellular functioning elicited by insulin. To our knowledge this would be the first demonstration of a dysfunction in the G-protein system caused by a pathological state. Such a lesion may account in part for changes in hormonal sensitivity observed in diabetes and may contribute to the expression of some of the many and various complications that characterize type I diabetes.

We thank Societe d'Etudes et de Recherches pour la Pharmacie et l'Agriculture (SERPA, France), the MRC of the UK, Scottish Home and Health Department and the California Metabolic Research Foundation for financial support. D.G. thanks the MRC for a research studentship.

Ectoenzymes control adenosine modulation of immunoisolated cholinergic synapses

P. J. Richardson, S. J. Brown, E. M. Bailyes & J. P. Luzio

Department of Clinical Biochemistry, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, UK

One of the most important inhibitory modulators of synaptic transmission in mammalian brain is adenosine. At some cholinergic terminals, adenosine is known to inhibit further release of acetylcholine. It is unclear whether adenosine is released directly at the synapse or whether ATP is co-released with transmitter and hydrolysed to adenosine in the synaptic cleft¹⁻⁴. Methods used in the past for isolating nerve terminals have not yielded homogeneous preparations, making it impossible to determine whether sufficient ATP or adenosine is released at specific synapses for inhibition of transmitter release to occur. Immunoaffinity purification techniques⁵⁻⁸ have recently permitted the preparation of homogeneous populations of cholinergic nerve terminals⁹, which release ATP upon stimulation¹⁰. We now report that in immunoisolated cholinergic nerve terminals from the striatum synaptic ectophosphohydrolases convert this ATP to adenosine, which inhibits further acetylcholine release, but this inhibitory effect is not seen in cortical cholinergic terminals lacking the complete ectophosphohydrolase pathway. Therefore the differing adenosine-mediated modulation in different brain areas is controlled by the presence and activity of synaptic ectophosphohydrolases.

Cholinergic nerve terminals were affinity purified from rat brain using a specific antibody against cholinergic neurons (sheep anti-(Chol-1) serum¹¹) and a monoclonal mouse anti-(sheep immunoglobulin G(IgG)) immunoabsorbent^{8,9} as previously described¹². Briefly, partially purified nerve terminals were preincubated with anti-(Chol-1) serum and those terminals coated with sheep IgG then bound to the solid phase bearing the monoclonal anti-(sheep IgG) antibody. These were then separated from the free terminals by centrifuging through 45% (v/v) Percoll. Such purified terminals are metabolically active⁹ and when viewed through an electron microscope can be seen to contain synaptic vesicles, mitochondria and adhering post-synaptic membrane. They also show Ca^{2+} -dependent release of acetylcholine and ATP on depolarization with 75 μ M veratridine. The ratio of acetylcholine/ATP (9:1) released, was close to that found in cholinergic synaptic vesicles derived from both rat striatum¹⁰ and *Torpedo* electric organ^{13,14}, consistent with the vesicular hypothesis of acetylcholine release. Analysis of the purines released on stimulation showed that the released ATP was rapidly degraded by a series of ectoenzymes (intrinsic plasma membrane enzymes with extracellularly directed active sites) including 5'-nucleotidase¹⁰. Such extracellular phosphohydrolase pathways have been found in a number of different cell systems¹⁵ including cholinergic terminals derived from the *Torpedo* electric organ^{16,17}. At the striatal cholinergic nerve terminal the ectoenzyme 5'-nucleotidase was found to be the rate limiting step in the production of adenosine from released ATP¹⁰.

To investigate acetylcholine release from immunoaffinity-purified cholinergic nerve terminals, the terminals were incubated with ³H-choline and placed in superfusion chambers. In each experiment the amount of released ³H-acetylcholine was determined after two 2-min periods of veratridine stimulation (S_1 and S_2 , see Fig. 1 legend). In control experiments, ³H-acetylcholine release in S_2 was always less than in S_1 . The S_2/S_1 ratios were compared under different conditions and it was found that, in agreement with previous observations on heterogeneous tissue preparations, addition of acetylcholine had

Received 15 January; accepted 2 April 1987.

- Houslay, M. D. *Trends biochem. Sci.* **9**, 39-40 (1983).
- Houslay, M. D. *Molec. Asp. cell Reg.* **4**, 279-333 (1985).
- Heyworth, C. M., Whetton, A. D., Wong, S., Martin, B. R. & Houslay, M. D. *Biochem. J.* **228**, 593-603 (1985).
- O'Brien, R. M., Houslay, M. D., Milligan, G. & Siddle, K. *FEBS Lett.* **212**, 281-288 (1987).
- Wilson, R. M. & Reeves, W. G. *Recent Adv. Diabetes* **2**, 127-138 (1986).
- Martin, J. M., Hunter, D. D. & Nathanson, N. M. *Biochemistry* **24**, 7521-7525 (1985).
- Yatani, A., Codina, J., Brown, A. M. & Birnbaumer, L. *Science* **235**, 207-209 (1987).
- Heyworth, C. M., Hanski, E. & Houslay, M. D. *Biochem. J.* **222**, 189-194 (1984).
- Itoh, H., Okajima, F. & Ui, M. *J. biol. Chem.* **259**, 15464-15473 (1984).
- Hildebrandt, J. D., Hanoune, J. & Birnbaumer, L. *J. biol. Chem.* **257**, 14723-14725 (1982).
- Katada, T. & Ui, M. *J. biol. Chem.* **257**, 7210-7216 (1982).
- Pushpendran, C. K., Corvera, S. & Garcia-Sainz, J. A. *FEBS Lett.* **160**, 198-202 (1983).
- Itoh, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3776-3780 (1986).
- Nukada, T. *et al. FEBS Lett.* **197**, 305-310 (1986).
- Heyworth, C. M., Wallace, A. V. & Houslay, M. D. *Biochem. J.* **214**, 99-110 (1983).
- Houslay, M. D. & Elliott, K. R. F. *FEBS Lett.* **104**, 359-363 (1979).
- Heyworth, C. M., Grey, A.-M., Wilson, S. R., Hanski, E. & Houslay, M. D. *Biochem. J.* **235**, 145-149 (1986).
- Sauerheber, R. D., Kuhn, C. E. & Hyslop, P. A. *Diabetes* **33**, 258-265 (1984).
- Milligan, G., Gierschik, P., Spiegel, A. M. & Klee, W. A. *FEBS Lett.* **195**, 225-230 (1986).
- Tanabe, T. *et al. Nature* **315**, 242-245 (1985).
- Yatsunami, K. & Khorana, H. G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4316-4320 (1985).
- Falloon, J. *et al. FEBS Lett.* **209**, 352-356 (1986).
- Milligan, G. & Klee, W. J. *J. biol. Chem.* **260**, 2057-2063 (1985).

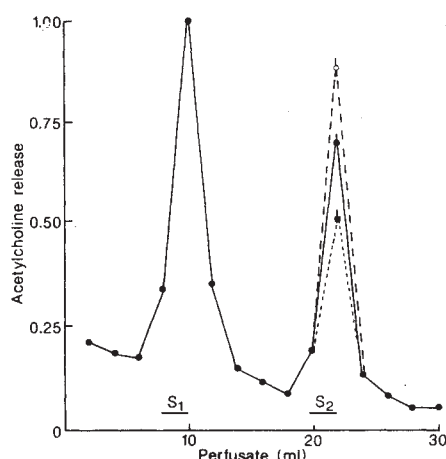


Fig. 1 Effect of ecto-5'-nucleotidase activity on acetylcholine release from affinity-purified nerve terminals. The amount of ^3H -acetylcholine released after each stimulation was calculated by subtracting the average basal release and the results normalized to a peak S_1 release of 1.0. The control (●) ratio was 0.60 ± 0.02 , whereas inhibition of 5'-nucleotidase (○) resulted in an increase in the S_2/S_1 ratio of 0.12 ± 0.03 . An additional 20 mU of 5'-nucleotidase (■) reduced the S_2/S_1 ratio by 0.20 ± 0.03 . Control and experimental perfusions were run on the same batch of terminals on the same day. Significance of the differences from the control ratio was determined by paired *t*-tests, and the probability $P < 0.05$ in both cases. Duration of stimulus is represented by the bars.

Methods. Affinity-purified striatal cholinergic nerve terminals⁹ were preincubated for 5 min at 37°C in Krebs-Ringer-HEPES, then with ^3H -choline ($10 \mu\text{Ci ml}^{-1}$, $1.2 \mu\text{M}$) for 5 min, washed and in some cases an immunoglobulin fraction corresponding to a 1:50 dilution of rabbit anti-rat 5'-nucleotidase (purified to homogeneity using a monoclonal antibody³¹). The terminals were then placed in a superfusion chamber (0.2 ml volume) over a glass fibre filter (Whatman GF/C). The terminals were perfused at 37°C with Krebs-Ringer-HEPES at a rate of 1 ml min^{-1} for 12 min in the presence of 1 mU ml^{-1} adenosine deaminase (to remove endogenous extraterminal adenosine). The terminals were then stimulated with $7.5 \times 10^{-5} \text{ M}$ veratridine for 2 min (S_1). After a further 10-min perfusion the terminals were again stimulated with veratridine. Some 5'-nucleotidase (20 mU), affinity purified and bound to a cellulose immunoadsorbent bearing a non-inhibitory mouse monoclonal antibody against 5'-nucleotidase, was included in the perfusion chambers in some cases.

no effect¹⁸ but adenosine inhibited acetylcholine release from striatal terminals^{10,19} (Table 1). We have already shown that this adenosine inhibition was sensitive to methylxanthines and adenosine deaminase and potentiated by the adenosine uptake inhibitor dipyridamole¹⁰. The adenosine analogue 2-chloroadenosine also inhibited acetylcholine release, both in the presence (35% inhibition, Table 1) and absence (38% inhibition) of dipyridamole. It therefore appears that this analogue is not a good substrate of the adenosine uptake mechanism, as has been previously suggested²⁰. Measurement of the amount of ATP released on stimulation of the cholinergic nerve terminals by veratridine suggested that ATP could reach a concentration of between 10^{-4} M and $3 \times 10^{-3} \text{ M}$ in the synaptic cleft depending on the assumptions made as to the dimensions of the cleft¹⁰. Perfusion of the terminals with 10^{-3} M ATP resulted in 90% hydrolysis of the ATP and 36% inhibition of acetylcholine release, whereas the non-hydrolysable analogue $\alpha\beta$ -methylene-ATP had no effect (Table 1). Both theophylline and 90% inhibition of 5'-nucleotidase (by a specific polyclonal antibody²¹) prevented any effect of ATP on acetylcholine release (Table 1). Therefore the inhibition by ATP was effected by an adenosine receptor after hydrolysis of ATP.

To investigate whether endogenous ATP release also inhibits acetylcholine release, 5'-nucleotidase was inhibited by the

Table 1 Modulation of acetylcholine release by purines

	Change (%)
Adenosine (10^{-5} M)	-26.3 ± 8.7 (3)
Adenosine (10^{-9} M)	NS (4)
Adenosine (10^{-9} M) plus dipyridamole	-29.1 ± 2.7 (24)
Adenosine (10^{-9} M) plus dipyridamole and theophylline	NS (9)
2-Chloroadenosine (10^{-10} M) plus dipyridamole	-35.3 ± 1.9 (14)
2-Chloroadenosine (10^{-10} M) plus dipyridamole and IBMX	NS (3)
Dipyridamole	-30.2 ± 4.3 (9)
Dipyridamole plus adenosine deaminase (1 mU ml^{-1})	NS (3)
Dipyridamole plus theophylline	NS (4)
ATP (10^{-3} M)	-36.2 ± 3.4 (20)
ATP (10^{-3} M) plus theophylline	NS (6)
ATP (10^{-3} M) plus anti-(5'-nucleotidase)	NS (4)
ATP (10^{-3} M) plus 5'-nucleotidase	-44.2 ± 3.2 (8)
ATP (10^{-3} M) plus 5'-nucleotidase and theophylline	-20.1 ± 4.1 (4)
$\alpha\beta$ -Methylene-ATP (10^{-6} M)	NS (2)
5'-Nucleotidase	-20.2 ± 3.0 (8)
Inhibitory anti-(5'-nucleotidase)	$+12.4 \pm 2.7$ (13)

The release of acetylcholine from affinity-purified striatal cholinergic nerve terminals was assessed as described in the legend to Fig. 2. The control condition contained adenosine deaminase (1 mU ml^{-1}) up to the end of S_1 . In those experiments investigating the effects of added adenosine and 2-chloroadenosine, these modulators were added after S_1 . In experiments examining the effects of endogenous adenosine and ATP release, adenosine deaminase was omitted from S_1 , unless stated otherwise. The inhibitory polyclonal anti-(5'-nucleotidase) antibodies (immunoglobulin fraction) were added at a final dilution of 1:40. IBMX (isobutylmethylxanthine), used at 10^{-6} M ; theophylline and dipyridamole were present at a concentration of 10^{-5} M . Purified 5'-nucleotidase was affinity purified on a non-inhibitory monoclonal antibody. Where stated, 20 mU of enzyme were added to the purified terminals. The results are expressed as % changes in the S_2/S_1 ratio with respect to the control condition and are means $\pm$ s.e.m. (with the number of experiments in parentheses) or means $\pm$ range (of 2 experiments). NS, not significant, $P > 0.05$.

specific antibody and the S_2/S_1 ratios compared with a control condition. It can be seen from Fig. 1 that inhibition of ATP breakdown to adenosine resulted in an increase in the size of S_2 . In contrast, S_2 was reduced by adding affinity-purified 5'-nucleotidase. These observations showed that in immunopurified striatal cholinergic nerve terminals, ectophosphohydrolase-mediated breakdown of released ATP is the major feedback autoinhibitory mechanism controlling acetylcholine release. Increases in the activity of the ectophosphohydrolase pathway potentiated the autoinhibition.

Similar experiments were performed with cholinergic nerve terminals derived from the cerebral cortex, in which acetylcholine has been shown to be a feedback inhibitor of its own release in a heterogeneous tissue preparation¹⁸. In contrast to the data from striatal cholinergic nerve terminals (Table 1) no inhibition of veratridine-stimulated acetylcholine release from cerebral cortex terminals was observed with extracellular adenosine (in agreement with previous observations¹⁹), 2-chloroadenosine and ATP (data not shown). Figure 2 shows that the cortical terminals, in contrast to striatal terminals, can produce only low amounts of nucleosides (2% compared to 25% of total added) from added extracellular ATP owing to the low activities of both ecto-5'-nucleotidase and ecto-(ADP)ase. The cortical terminals, although capable of ATP inactivation, neither hydrolyse ADP to adenosine nor do they express presynaptic adenosine receptors.

That the breakdown of presynaptically released ATP results in autoinhibition in rat brain striatal cholinergic nerve terminals confirms a physiological function for ectoenzymes (and ectophosphohydrolases in particular) in the synaptic cleft. It is

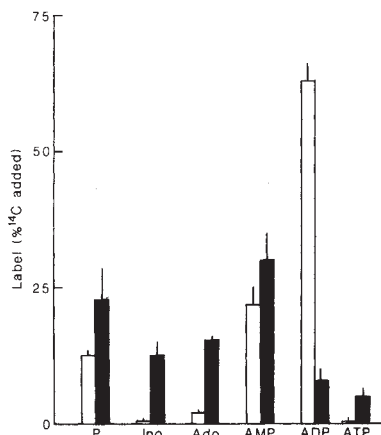


Fig. 2 Ectophosphohydrolase-mediated hydrolysis of ATP by cortical cholinergic nerve terminals. Labelled ^{14}C -ATP (~ 15 pmol) was added to affinity-purified nerve terminals in the presence of 7.5×10^{-5} M veratridine. After 2 min at 37°C , the terminal suspensions were centrifuged and the metabolites present in the supernatant separated by HPLC and their radioactivity (as well as the label in the pellet) determined¹⁰. The results are expressed as a percentage of the total label added and are means $\pm$ s.e.m. from three experiments (solid bars, striatum), or seven experiments (open bars, cortex). P, ^{14}C present in pellet; Ino, inosine; Ado, adenosine.

often difficult to study the physiological function of ectoenzymes because the source of the substrate is unknown²². Affinity purification of cholinergic nerve terminals has allowed us to show that the substrate for ectophosphohydrolases at cholinergic synapses is presynaptic vesicular ATP. Because 5'-nucleotidase is not detectable (by antibody and complement-mediated lysis tests²³) on the presynaptic membrane of cholinergic nerve terminals, it appears that the feedback autoinhibition of transmitter release is determined both by the presynaptic terminal (which releases ATP) and the postsynaptic membrane bearing ectophosphohydrolases. It is clear that greater feedback inhibition can be obtained by increasing the amount of 5'-nucleotidase, suggesting that further inhibition of these synapses could be exerted by either adjacent glial cells (which are known to have ecto-5'-nucleotidase activity^{24,25}), or by increases in the synthesis and expression of postsynaptic 5'-nucleotidase. Because ATP is co-packaged with a number of neurotransmitters (such as noradrenaline²⁶, serotonin²⁷ and acetylcholine) whose release can be inhibited by adenosine^{28,29}, it is possible that synaptic ectophosphohydrolases have a general function in the autoinhibition of neurotransmission. Differential expression and activity of these ectoenzymes may be a mechanism by which long-term synaptic plasticity occurs. It is also conceivable that ectoenzymes are responsible for controlling the concentrations of other neuromodulators in the brain, for example, the ecto-endopeptidases that cause neuropeptide inactivation³⁰.

Cholinergic synapse autoinhibition in the striatum, where there is adenosine-mediated but not acetylcholine-mediated feedback, contrasts with that in the cerebral cortex. These cortical terminals are inhibited by acetylcholine but not by adenosine; nor is there ectophosphohydrolase activity present capable of degrading ATP to adenosine. The results suggest that different rat brain cholinergic neurons have two distinct feedback control mechanisms, one determined by presynaptic adenosine receptors and the activity of ectophosphohydrolases on adjacent cells (in the striatum), and one determined by presynaptic inhibitor acetylcholine receptors (in the cortex). It is probable that the developmental decision as to which autoinhibitory mechanism is operational is related to the function of the neurons.

This work was supported by grants from the Muscular Dystrophy Group and the MRC. S.J.B. is the recipient of an MRC Training Award, and we thank A. C. Newby and A. Fine for many helpful discussions.

Received 10 February; accepted 2 April 1987.

- De Mey, J., Burnstock, G. & Vanhoutte, P. M. *Eur. J. Pharmac.* **55**, 401-405 (1979).
- Silinsky, E. M. *Br. J. Pharmac.* **71**, 191-194 (1980).
- Stone, T. W. *J. Physiol., Lond.* **335**, 591-608 (1983).
- Moody, C. J. & Burnstock, G. *Eur. J. Pharmac.* **77**, 1-9 (1982).
- Luzio, J. P., Newby, A. C. & Hales, C. N. *Biochem. J.* **154**, 11-21 (1976).
- Ito, A. & Palade, G. E. *J. Cell Biol.* **79**, 590-597 (1978).
- Gruenberg, J. & Howell, K. E. *EMBO J.* **5**, 3091-3101 (1986).
- Richardson, P. J. & Luzio, J. P. *Appl. Biochem. Biotechnol.* **13**, 133-145 (1986).
- Richardson, P. J., Siddle, K. & Luzio, J. P. *Biochem. J.* **219**, 647-654 (1984).
- Richardson, P. J. & Brown, S. J. *J. Neurochem.* **48**, 622-630 (1987).
- Richardson, P. J., Walker, J. H., Jones, R. T. & Whittaker, V. P. *J. Neurochem.* **38**, 1605-1614 (1982).
- Richardson, P. J. *J. Neurochem.* **46**, 1251-1255 (1986).
- Tashiro, T. & Stadler, H. *Eur. J. Biochem.* **90**, 479-487 (1978).
- Zimmermann, H. in *Co-transmission* (ed. Cuello, A. C.) 243-259 (Macmillan, London, 1982).
- Pearson, J. D., Carleton, J. S. & Gordon, J. L. *Biochem. J.* **190**, 421-429 (1980).
- Zimmermann, H., Dowdall, M. J. & Lane, D. A. *Neuroscience* **4**, 979-993 (1979).
- Zimmermann, H., Grondal, E. J. M. & Kellett, F. in *Cell Biology of Ecto-enzymes* (eds Kreutzberg, G. W., Reddington, M. & Zimmermann, H.) 35-48 (Springer, Berlin, 1986).
- Marchi, M., Paudice, P., Caviglia, A. & Raiteri, M. *Eur. J. Pharmac.* **91**, 63-68 (1983).
- Pedata, F., Antonelli, T., Lambertini, L., Beani, L. & Pepeu, G. *Neuropharmac.* **22**, 609-614 (1983).
- Daly, J. W. *J. med. Chem.* **25**, 197-207 (1982).
- Stanley, K. K., Burke, P., Pitt, P., Siddle, K. & Luzio, J. P. *Expl Cell Res.* **144**, 39-46 (1983).
- Stanley, K. K., Newby, A. C. & Luzio, J. P. *Trends Biochem. Sci.* **7**, 145-147 (1982).
- Richardson, P. J. *J. Neurochem.* **41**, 640-648 (1983).
- Kreutzberg, G. W., Barron, K. D. & Schubert, P. *Brain Res.* **158**, 247-257 (1978).
- Kreutzberg, G. W., Heymann, D. & Reddington, M. in *Cellular Biology of Ecto-enzymes* (eds Kreutzberg, G. W., Reddington, M. & Zimmermann, H.) 147-164 (Springer, Berlin, 1986).
- Lagerkrantz, H. *Acta physiol. scand. Suppl.* **366**, 1-44 (1971).
- Da Prada, M. & Pletscher, A. *Br. J. Pharmac.* **34**, 591-597 (1968).
- Harms, H. H., Wardeh, G. & Mulder, A. H. *Eur. J. Pharmac.* **49**, 305-308 (1978).
- Harms, H. H., Wardeh, G. & Mulder, A. H. *Neuropharmacology* **18**, 577-580 (1979).
- Relton, J. M., Fee, N. S., Matsas, R., Turner, A. J. & Kenny, A. J. *Biochem. J.* **215**, 519-523 (1983).
- Baillyes, E. M., Newby, A. C., Siddle, K. & Luzio, J. P. *Biochem. J.* **203**, 245-251 (1982).

Cell lineages generating axial muscle in the zebrafish embryo

Charles B. Kimmel & Rachel M. Warga

Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403, USA

Cell lineage may contribute to determining the numbers, positions and types of cells formed during embryogenesis. *In vitro* clonal analyses show that vertebrate cells can autonomously maintain lineage commitments to single fates¹⁻⁵ and that terminal development may include an invariant sequence of cell divisions^{6,7}. In addition, *in vivo* studies with *Xenopus*⁸ led to the proposal that clonal restrictions to spatial 'compartmental' domains arise during early development, analogous to what is observed in insects^{9,10}. In the zebrafish, individual gastrula cells generate clones of progeny that are confined within single tissues¹¹, but spatial restrictions have not been described. We now have examined the *in vivo* terminal cell lineages of zebrafish axial muscles. We obtained no evidence either for strict developmental regulation of division pattern or for spatial compartmentation within muscle lineages.

In somitic muscle the formation of segment-long striated cells provides a convenient marker for terminal cell differentiation (Fig. 1a), and the rigid geometry of the myotome allows unambiguous regional assignments for individual cells (Fig. 1b). In many vertebrates muscle fibres arise by fusions of mononucleate cells. This would complicate cell lineage analysis if, for example, a cell from one lineage fused with a cell from another. At about 30 hours after fertilization (denoted 30 h) more than 70% of muscle fibres in the zebrafish are mononucleate, as revealed by counting the number of nuclei in individual muscle fibres labelled with the lineage tracer horseradish peroxidase (Fig. 1c-e).

Fig. 1 Differentiated muscle fibres in the 24–30 h zebrafish embryo. *a*, Photomicrograph of the myotomes of the left side of a live embryo at 30 h, using Nomarski optics. Elongated muscle fibres, in which cell nuclei and cross-striations are evident, stretch between the myotomal borders. The horizontal myoseptum runs from left to right near the upper part of the photograph. *b*, Individual fibres, filled with fluorescein isothiocyanatedextran after injection¹⁵ of their precursor blastomere, photographed from the left side of a live embryo by epifluorescence. The three upper fibres (which appear shorter because they are out of focus) are in the dorsal parts of two adjacent myotomes and the three lower fibres are in the ventral parts of these same two myotomes. *c*, *d*, Single muscle fibres, labelled with horseradish peroxidase, containing one (*c*) and two (*d*) nuclei. *e*, Frequency histogram showing the numbers of nuclei in 290 single peroxidase-labelled muscle fibres that were distinct from other labelled fibres and were entirely within a single section. Material in *c–e* is from a library of sections through embryos in which single blastomeres (at a variety of stages) had been injected with horseradish peroxidase lineage tracer¹⁵. The embryos were later fixed and sectioned longitudinally at 50 μm and peroxidase activity revealed histochemically¹⁶. The micrographs (*a–d*) are all $\times 300$; scale bar (*d*), 25 μm .

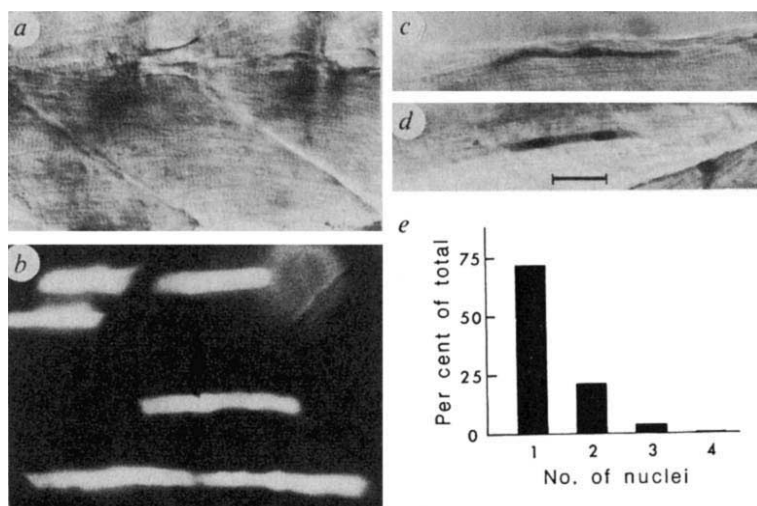
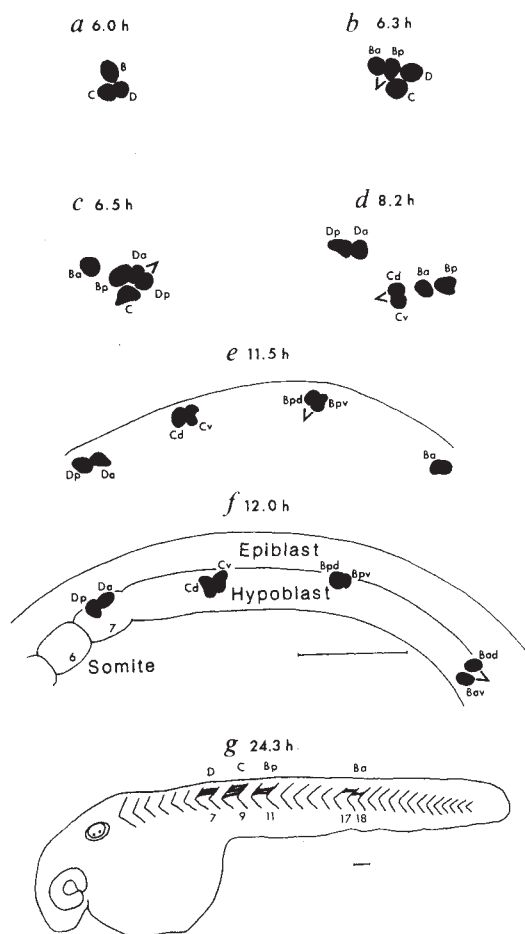


Fig. 2 Lineages of clonally related muscle cells in a zebrafish embryo. The same lineages are shown diagrammatically in Fig. 3a. A blastomere at the approximately 2,000-cell stage was injected with rhodamine-dextran, and in the gastrula at 6 h one or two enveloping layer cells (see ref. 11), and four deep cells were present. The deep cells were named A–D according to their distances from the animal pole, and cells B, C and D (*a*) eventually generated muscle fibres; the fourth deep cell (not shown) was in prospective neural ectoderm. Using a silicon-intensified-target video camera to amplify fluorescence, we repeatedly recorded the cell lineages until 13 h, using video tape. The figure shows tracings, all at the same orientation (dorsal up, anterior to the left), from the face of the video monitor at selected times indicated. Tracings *a–f* are at the same magnification and *g* is reduced $\sim$ sevenfold. Scale bars (*f*, *g*), 50 μm . By 12 h (*f*) the epiblast and hypoblast layers of deep cells were well delineated, and the boundary between them is shown just above the cells. Seven somites had formed, and the last two are numbered. Once labelled cells were observed within a formed somite they were never observed to migrate out of that somite, which facilitate their subsequent identification. The relative positions of siblings shortly after their births are coded by their names: a, anterior; p, posterior; d, dorsal; v, ventral. Note that between 6.3 h and 11.5 h there are inversions in the anterior-posterior positions of the cells. This change appears to result from active cell movements that will be described elsewhere.



Thus, as in *Xenopus*¹², most of these cells differentiate by elongation without cell fusion. This simplifies analysis of their lineages.

To generate marked cells that we could follow during normal *in vivo* development, we injected a fluorescent lineage tracer dye into single blastomeres in embryos at midblastula stage (after the tenth or eleventh cleavage, at 3–3.5 h). Beginning at the early gastrula stage, when the blastoderm covered about half the yolk cell (~ 6 h), we watched labelled cells that were descended from the injected cell (Fig. 2a), recording divisions

and movements of some cells in each clone¹¹. We generally did not keep track of lineages of labelled cells in the enveloping layer on the surface of the gastrula, that later would form clones restricted to the peridermis, or surface layer, of the embryo¹¹. Furthermore, in some experiments we had to abandon following the lineages of some of the deep cells, because of uncertainties about particular lineage assignments. Observations were continued at intervals of about 15 min until several somites had formed (12–15 h; Fig. 2b–f). At this time, clonal lineages that

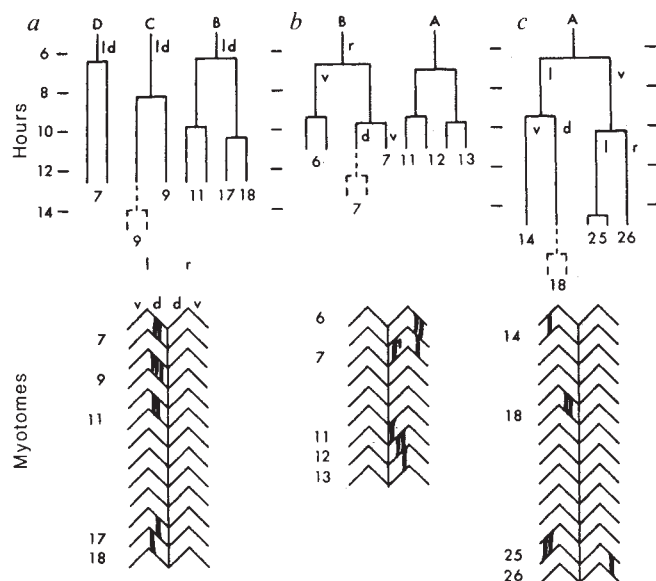


Fig. 3 Patterns of cell lineages of zebrafish muscle cells. The data in *a* were obtained from the embryo shown in Fig. 2, and in *b* and *c* from two other embryos, by observing the cells as they developed (see text and Fig. 2 legend). Only lineages that eventually formed muscle are shown; there were other labelled deep cells in each embryo (1 in *a*, 2 in *b*, and 11 in *c*) in addition to labelled enveloping layer cells. In one case in each embryo (see text) extra cell divisions must have occurred, after the times of observations, and these unobserved divisions are indicated by dashed lines. In all but one of the cases shown, each branch in each lineage diagram ends with a single muscle fibre, present in the myotome numbered below the branch. In the one exceptional case the branch ends with an undifferentiated cell that may not have been postmitotic (in myotome 7 of the embryo in *b*) and that lay adjacent to a labelled muscle fibre (the cell's presumed sibling). The letters along the branches indicate that all descendants of that branch were restricted to the same region of the segment; d, dorsal; v, ventral; l, left; r, right. Lower drawings, spatial distributions of the muscle fibres. The presentations are dorsal diagrammatic views, with the midline of the embryo in the centre of each drawing, and the left and right and the dorsal and ventral regions of the myotomes indicated.

would form muscle cells could be identified by the presence of the labelled cells in the paraxial hypoblast (Fig. 2f). The embryos then were incubated without further observation until 24–30 h when the cells had developed into striated muscle fibres and their positions could be charted (Fig. 2g).

Data for the embryo in Fig. 2 are shown in diagrammatic form in Fig. 3a. Progeny of cells B, C and D each eventually generated muscle fibres. Five cell divisions were watched, and by 12 h eight cells were present (upper part of Fig. 3a). We observed one cell pair (the D siblings) become sequestered within somite seven and another pair (the C siblings) become sequestered within somite nine. At 24.3 h there were a total of nine muscle fibres, distributed as expected from the previous observation (lower part of Fig. 3a). All the fibres were present in the dorsal parts of myotomes on the left side of the embryo. The extra fibre was in segment 9.

These data suggested that we had watched all but one of the cell divisions that generated these muscle fibres and that one of the C siblings (Cd, dorsal or Cv, ventral) divided once more during the period between 12 and 24 h when the embryo was left unobserved. Because cell fusions are infrequent (Fig. 1e), it is unlikely that extra divisions coupled with cell fusions occurred. To describe the relationship between cell replication and terminal differentiation, we define a 'birthday' division to be one in which one or both of the daughter cells differentiate into postmitotic muscle fibres. Many cells had birthday divisions

between about 6 and 15 h. In the selected cases in Fig. 3, most birthday divisions occurred between 8 and 10 h, but some cells, especially those located in caudal segments, underwent an additional division after 10 h. In all, we could make clear assignments of the fates of cells arising from 35 divisions, and 26 of these were birthday divisions (generating 48 muscle fibres in 15 clones from 7 embryos).

In two of the clones in Fig. 3 (one each in *a* and *b*) two rounds of divisions generated four muscle fibres; the lineages are symmetrical. Symmetrical lineages were proposed by Quinn *et al.*⁶ to explain why the final number of cells found in chick muscle cell clones *in vitro* was frequently observed to be an integral exponent of two, as if there were a mechanism for counting cell divisions. However, we observed in other cases (overall, 5 out of 12 subclones that included two or more rounds of divisions) that only one of the daughter cells arising from a birthday division generated a postmitotic fibre; the other daughter divided again, so that the clone generated three (Fig. 3a), five (Fig. 3b) or six (Fig. 3c) cells. Thus our data provide little evidence for a division-counting mechanism that regulates these muscle cell lineages *in vivo*. A recent *in vitro* clonal analysis, where chick muscle cell lineages were traced by time-lapse cinematography, also revealed many exceptions to symmetrical lineages¹³.

The symmetrical lineages we did observe (in 7 of 12 clones) could be explained if sibling cells tend to divide synchronously before they differentiate. The cell cycle length determined in 38 cases in which we watched two successive divisions averaged 3.6 ± 1.0 ($\pm$ s.d.) hours. Among these cases were 16 pairs of siblings. Their divisions were very synchronous, within an average interval of 0.8 ± 0.6 hours. Because birthday divisions generally were hours ahead of overt differentiation (that is, cell elongation) it seems likely that this tendency to divide synchronously, and not division counting, is the explanation for the symmetrical lineages we observed. Konigsberg and Pfister found that sibling myogenic cells also divide synchronously *in vitro*¹³.

In *Xenopus*, blastomeres of 512-cell blastulae were reported to generate clones in which muscle cells are confined either to dorsal (epaxial) or to ventral (hypaxial) domains of the myotome¹⁴. We charted muscle fibre positions (as in the lower diagrams in Fig. 3) in clones derived from blastomeres injected at or after the 512-cell stage in 36 embryos, and observed no strict dorsal-ventral, segmental, or even left-right compartmentation of myotomes. The examples in Fig. 3 show that separations of cells across these putative compartment boundaries can occur at very late stages in the lineage: in the embryo in Fig. 3a, sibling muscle fibres were separated by the boundary between myotomes 17 and 18. In Fig. 3b a birthday division generated cells that became separated into the dorsal and ventral region of the same myotome (no. 7). In Fig. 3c there are many separations, including dorsal-ventral and left-right separations after birthday divisions. These results clearly rule out strict determination of muscle fibre position by cell lineage in the zebrafish.

Overall, there was considerable scattering of clonally related cells that generated muscle fibres, especially in tail segments (posterior to myotome 15). In fact all the observed crossings between left and right myotomes occurred in tail segments. This difference in the clonal origins of trunk and tail myotomes may be underlain by morphogenetic differences. The cells that will form trunk myotomes gather into the paraxial regions before the end of gastrulation (before the yolk cell is completely enveloped by the blastoderm). The tail myotomes arise from a 'tail bud' of cells that forms by cell convergence after gastrulation is completed (C.B.K. and R.M.W., unpublished observations). In contrast to these muscle lineages, neural lineages of the zebrafish give rise to cells that cross the midline within the central nervous system at any axial level¹¹. This suggests that at least some lineage patterning processes that operate during

early development are not global, but work differently for different tissues.

We thank Judith Eisen, Stephen Hauschka, Walter Metcalfe, and Monte Westerfield for discussions and for review of earlier versions of this paper. Irwin Konigsberg provided us with his manuscript before its publication. Harrison Howard and Reida Kimmel helped with the photography. The research was supported by a grant from NSF.

Received 24 December 1986; accepted 3 March 1987.

1. Konigsberg, I. R. *Science* **140**, 1273-1284 (1963).
2. Rutz, R. & Hauschka, S. *Dev. Biol.* **91**, 103-110 (1982).
3. Womble, M. D. & Bonner, P. H. *J. Embryol. exp. Morph.* **58**, 119-130 (1980).
4. Coon, H. C. *Proc. natn. Acad. Sci. U.S.A.* **55**, 66-73 (1966).
5. Cahn, R. D. & Cahn, M. B. *Proc. natn. Acad. Sci. U.S.A.* **55**, 106-114 (1966).
6. Quinn, L. S., Holtzer, H. & Nameroff, M. *Nature* **313**, 692-694 (1985).
7. Temple, S. & Rafi, M. C. *Cell* **44**, 773-779 (1986).
8. Jacobson, M. J. *Neurosci.* **3**, 1019-1038 (1983).
9. Garcia-Bellido, A., Ripoll, P. & Morata, G. *Nature new Biol.* **245**, 251-253 (1973).
10. Crick, F. H. C. & Lawrence, P. A. *Science* **189**, 340-347 (1975).
11. Kimmel, C. B. & Warga, R. M. *Science* **231**, 365-368 (1986).
12. Muntz, L. J. *Embryol. exp. Morph.* **33**, 757-774 (1975).
13. Konigsberg, I. R. & Pfister, K. K. *Expl Cell Res.* **167**, 63-74 (1986).
14. Moody, S. A. & Jacobson, M. J. *Neurosci.* **3**, 1670-1682 (1983).
15. Kimmel, C. B. & Law, R. D. *Dev. Biol.* **108**, 94-101 (1985).
16. Metcalfe, W. K., Kimmel, C. B. & Schabach, E. J. *comp. Neurol.* **233**, 377-389 (1985).

Low-density lipoproteins inhibit endothelium-dependent relaxation in rabbit aorta

H. E. Andrews*, K. R. Bruckdorfer*, R. C. Dunn* & M. Jacobs†

* Department of Biochemistry and † Department of Pharmacology, Royal Free Hospital School of Medicine, London NW3 2PF, UK

The vascular endothelium, in response to pulsatile flow and vasoactive agents including acetylcholine^{1,2}, secretes the endothelium-derived relaxing factor (EDRF), a substance which regulates vascular tone. Recent interest in EDRF has focused on its possible dysfunction in atherosclerosis¹. In animal models of the disease, endothelium-dependent relaxation is markedly reduced^{3,4}. The continuous exposure of the endothelium in hyperlipidaemia to high concentrations of low-density lipoprotein (LDL), a known atherogenic risk factor, may explain this dysfunction. Here, we demonstrate that pathophysiological concentrations of LDL directly inhibit endothelium-dependent relaxation. Chemically modified LDL, in contrast, is inactive, implying that the inhibition is through a receptor-dependent mechanism.

Human or rabbit LDL (relative density 1.006-1.063 g ml⁻¹), prepared by the method of Chung⁵, was used throughout. Descending thoracic aortae were removed from 6-month-old Japanese White rabbits (2-3 kg), and trimmed free of adhering fat and connective tissue. Transverse rings 2-mm wide were cut and then mounted under 2g resting tension for isometric force measurements in oxygenated Krebs-bicarbonate buffer containing 0.03 mM EDTA at 37 °C, as previously described². Tissues were equilibrated for 90 min and dose-response curves determined with serotonin or noradrenaline. The effect of cumulative concentrations of relaxing agents was determined before and after treatment with LDL in tissues which were precontracted with a dose of serotonin or noradrenaline that gave 75% of the maximal contraction.

Figure 1a shows a typical recording of the dose-dependent relaxation evoked by acetylcholine (ACh) in an aortic ring precontracted with serotonin. The relaxations were unaffected by 10 μ M indomethacin and were therefore not due to the release of cyclooxygenase products (not shown). Furthermore, the removal of the endothelium by rubbing abolished the relaxation induced by ACh (Fig. 1b). The relaxation therefore has the

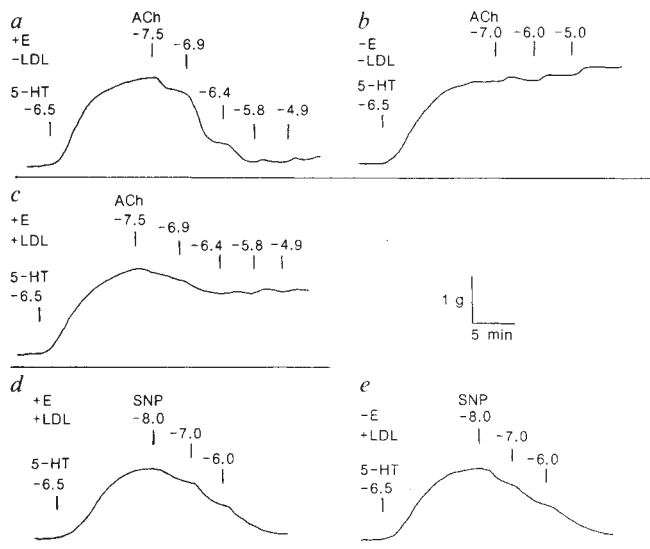


Fig. 1 The effect of LDL on endothelium-dependent and independent relaxation in rabbit aortic rings. Intact rabbit aortic rings and rings denuded of endothelium, prepared by inserting a pair of forceps and gently rolling the ring on moistened filter paper, were mounted in a jacketed water bath containing Krebs solution bubbled with 95% O₂ + 5% CO₂ and maintained at 37 °C. Tissues were precontracted with a predetermined dose of serotonin (5HT) which gave 75% maximal tone and then relaxed by cumulative concentrations of ACh. After washing for 15 min and recovery for 30 min, the tissues were incubated for 30 min with 2 mg protein ml⁻¹ LDL or Tyrode's buffer. The wash-recovery cycle was repeated and the pre-contracted tissues retested for relaxation with ACh (a-c) or sodium nitroprusside (SNP) (d and e). The integrity of the endothelium was assessed by Ag staining¹⁴ and found to be preserved before and after LDL treatment. Recordings are shown from intact (+E) rings (a, c, d) and rings denuded endothelium (-E, b and e) after preincubation with Tyrode's buffer (a and b) or with LDL at 2 mg protein ml⁻¹ (c-e). Cumulative molar concentrations (numbers above marker bars) are expressed as log units. **Methods.** LDL was prepared by discontinuous density gradient ultracentrifugation in a Kontron vertical rotor (TFT 70.38) using a Centrikon T2080 ultracentrifuge⁵. To remove contaminating plasma proteins the LDL was recentrifuged twice for 10 h at 190,000g ($r=6.75$ cm) in buffer with density adjusted to 1.063 g ml⁻¹ in a conventional rotor. Buffers were sometimes degassed and contained 0.25 mM thiomersal and 0.3 mM EDTA to prevent oxidation. The purified LDL was finally dialysed for 24 h against three changes of 2 l Tyrode's buffer containing 0.3 mM EDTA before use. LDL prepared by this procedure was found not to produce significant oxidized LDL when tested by fluorescence measurements (excitation 350 nm, emission 420 nm) as described by Koller¹⁰. The final LDL preparation migrated as a single polypeptide, apolipoprotein B in SDS gel electrophoresis.

characteristics of an endothelium-dependent relaxation mediated by EDRF².

The effect of LDL treatment on endothelium-dependent relaxation induced by ACh was investigated. A typical inhibition of relaxation by LDL at a concentration of 2 mg protein per ml is shown in Fig. 1c. This inhibitory effect depended on the time of preincubation with LDL before the wash-recovery cycle. Incubations longer than 10 min were required for appreciable inhibition to occur. After 20 min, inhibition was 30% of the final level and by 30 min, near maximal inhibition was reached and was maintained even after 2 h with extensive washing. The maximal inhibition of relaxation evoked by ACh after preincubation with 2 mg protein per ml of LDL is shown in Table 1 and is not significantly different whether serotonin or noradrenaline is used to pre-contrast the preparations (Fig. 1a and c; Fig. 2a and b). Whereas LDL inhibited endothelium-dependent relaxation of the intact rings, the contraction in response to

Table 1 Comparison of the maximal inhibition of endothelium dependent relaxation by LDL and modified LDL in rabbit aortic rings

Agonist	Relaxing agent	Maximal relaxation before pretreatment (%)	Pretreatment (2 mg protein ml ⁻¹)	Maximal inhibition (%)
3.2 × 10 ⁻⁷ M serotonin	4 × 10 ⁻⁷ M ACh	76.1 ± 2.6	LDL	62.4 ± 11
	1.3 × 10 ⁻⁷ M A23187	79.3 ± 3.7	LDL	48.3 ± 5
			LDL	55.0 ± 6.0
3.2 × 10 ⁻⁷ M noradrenaline	4 × 10 ⁻⁷ M ACh	68.4 ± 6.2	CHD-LDL	15.1 ± 6.5
			Reversed CHD-LDL	50.0 ± 6.7
	10 ⁻⁴ M ATP	75.2 ± 3.7	Carbamylated LDL	5 ± 5.4

Aortic rings were precontracted with noradrenaline or serotonin and maximal relaxation determined with cumulative concentrations of an endothelium-dependent relaxant before and after treatment with LDL as in the legend to Figs 1 and 2. LDL was modified by treatment with cyclohexanedione, (CHD-LDL), by the method of Mahley *et al.*⁸. Reversal of the modification (reversed CHD-LDL) was performed by incubation with hydroxylamine⁸ and carbamylated LDL was prepared by incubation with potassium cyanate⁹. Modified LDL was dialysed against three changes of 2 L Tyrode's buffer containing 0.3 mM EDTA before use. Maximal percentage relaxations of NA contractions, before and after treatment with 2 mg protein ml⁻¹ of LDL were compared usually at the same relaxant concentration as shown, and maximal percentage inhibition calculated. Changes in maximal percentage relaxation before treatment and in paired controls after treatment with Tyrode's buffer were not significantly different. Maximal percentage relaxation and inhibition are expressed as mean ± s.e.m. units and are the mean of 4-5 observations. Statistical significance was tested using Student *t*-test, *P* < 0.05 was considered significant.

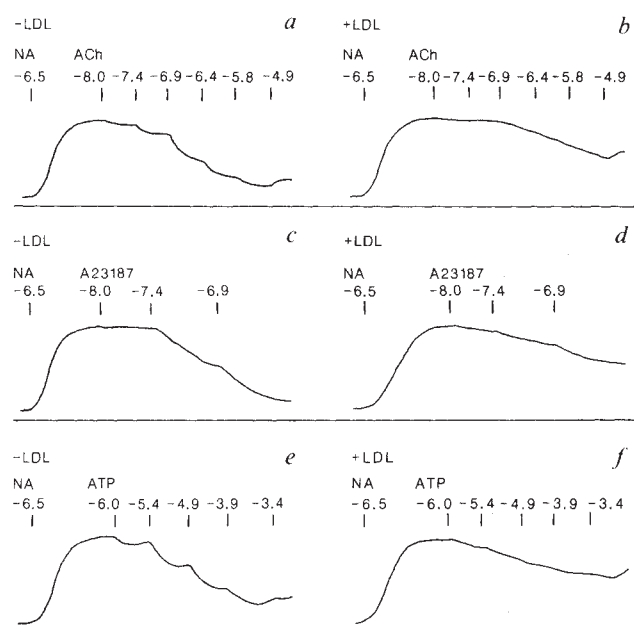


Fig. 2 Inhibition by LDL of relaxation evoked by ACh, A23187 and ATP in aortic rings precontracted with noradrenaline. Experiments were performed as in Fig. 1. Intact rabbit aortic rings containing endothelium were precontracted with noradrenaline and relaxed with cumulative concentrations of ACh (*a* and *b*), A26187 (*c* and *d*) or ATP (*e* and *f*). Recordings are from tissues preincubated with Tyrode's buffer (*a*, *c* and *e*) or a 2 mg protein ml⁻¹ LDL (*b*, *d* and *f*). Cumulative molar concentrations are expressed as log units above marker bars.

serotonin by intact or endothelium-denuded preparations was unaltered by exposure to LDL (Fig. 1*b*, *c* and *e*). Similarly sodium nitroprusside, whose action is known to be directly on vascular smooth muscle^{1,6}, relaxed intact or endothelium-denuded rings before and after treatment with LDL (Fig. 1*d* and *e*). Similar results were obtained if rabbit LDL was used instead of human LDL (not shown).

LDL also inhibited the relaxation evoked by the Ca²⁺ ionophore A23187 (Fig. 2*c* and *d*), and ATP (Fig. 2*e* and *f*) in intact aortic rings precontracted with noradrenaline. The former agent is endothelium-dependent in rabbit aorta^{1,6}. ATP, however, is known to have its relaxant action through the endothelium as well as a less potent direct effect in vascular smooth muscle, possibly by breakdown to AMP⁶. When endothelium-dependent relaxation by ATP was abolished by

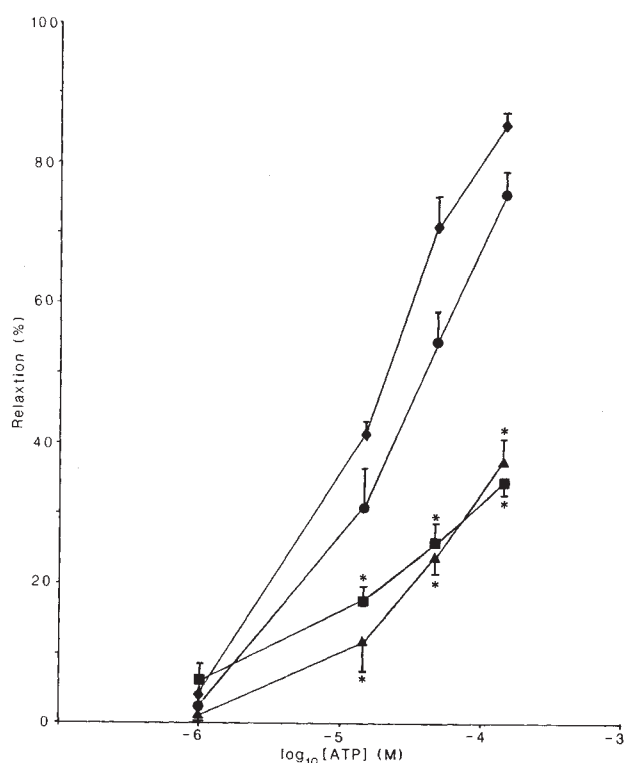


Fig. 3 Dose-response curves showing the effect of LDL on endothelium-dependent relaxation evoked by ATP in intact rabbit aortic rings precontracted with noradrenaline. Experiments were performed as in Fig. 1 using different concentrations of LDL with scale bar as Fig. 2. Recordings are shown from tissues relaxed with cumulative concentrations of ATP after treatment for 30 min with Tyrode's buffer (●) or LDL at a protein concentration of 0.5 mg ml⁻¹ (◆), 1 mg ml⁻¹ (▲) or 2 mg ml⁻¹ (■). Relaxations are expressed as the percentage of relaxation of noradrenaline contractions. Each point is the mean of 3-4 observations; vertical bars indicate the s.e.m. Significant changes in relaxation, shown by asterisks, were found only between control and rings treated with 1 or 2 mg ml⁻¹ LDL. Comparisons were made using the Student *t*-test where *P* < 0.05 was considered significant.

10 μM haemoglobin⁷, the remaining direct action was unaffected by LDL (not shown).

The optimum protein concentration for the inhibition was determined. Dose-response curves to cumulative concentrations of ATP were performed on intact rings pre-contracted with 0.3 μM serotonin after pre-incubation with different concentra-

tions of LDL for 30 min as shown in Fig. 3. At a low concentration of LDL (0.5 mg protein ml⁻¹), a small leftward shift in the dose response curve was observed, which was not significant. Higher concentrations of LDL caused a rightward shift and a decrease in relaxation. Maximal inhibition of relaxation by ATP was found at concentrations of 1–2 mg protein ml⁻¹ and similar concentrations were necessary to inhibit maximally endothelium-dependent relaxation by A23187 or ACh. At this LDL concentration, the maximal percentage inhibition of relaxation evoked by these vasodilators was not significantly different, as shown in Table 1.

Cyclohexanedione-LDL⁸ and carbamylated-LDL⁹, lipoproteins chemically modified in their lysines and arginines respectively, were used to investigate the specificity of the inhibition. The results (Table 1) show that the inhibition of ATP- and ACh-evoked relaxations were reduced by 65 and 92% respectively when cyclohexanedione or carbamylated-LDL were used instead of unmodified LDL. Most (91%) of the inhibitory effect was restored by reversal of the cyclohexanedione modification (Table 1)⁸.

These experiments indicate that the inhibition of endothelium-dependent relaxation is not due to non-specific effects such as by-products of LDL oxidation. Furthermore, when LDL was deliberately oxidized by Cu²⁺ as assessed by fluorescence¹⁰, no enhancement in the inhibition was noted. A rapid sequestration of EDRF by LDL as proposed for haemoglobin is also ruled out by the time dependence of the inhibition⁷. LDL does not inhibit sodium nitroprusside, a direct-acting vasodilator. This agent, like EDRF itself, stimulates guanylate cyclase activity in

the smooth muscle¹¹; this shows that the action of LDL is not at the level of this enzyme. The fact that chemically modified LDL, which is unable to bind to the high-affinity LDL-receptor¹², also fails to inhibit relaxation, implicates a receptor-dependent mechanism. The relatively high concentration of LDL required for inhibition may suggest that the endothelial receptors have a lower affinity for LDL than those found in fibroblasts and liver¹². The concentration for inhibition corresponds to the elevated plasma levels found in individuals with hyperlipidaemia¹³, who as a consequence of this dysfunction are expected to have increased risk of vascular spasm and thrombosis.

This work is supported by the MRC, British Heart Foundation and the Peter Samuel Royal Free Fund.

Received 28 January; accepted 6 April 1987.

1. Vanhoutte, P. M., Rubanyi, G. M., Miller, V. M. & Houston, D. S. *A. Rev. Physiol.* **48**, 307–320 (1986).
2. Furchgott, R. F. & Zawadzki, J. V. *Nature* **288**, 373–376 (1980).
3. Coene, M.-C. *et al. Br. J. Pharmac.* **85**, 267P (1985).
4. Ibengwe, J. K. & Suzuki, H. *Br. J. Pharmac.* **87**, 395–402 (1986).
5. Chung, B. H., Wilkinson, T., Geer, J. C. & Segrest, J. P. *J. Lipid Res.* **21**, 284–291 (1980).
6. Furchgott, R. F. *Circulation Res.* **53**, 557–573 (1983).
7. Martin, W., Villani, G. M., Jothianandan, D. & Furchgott, R. F. *J. Pharmac. exp. Ther.* **232**, 708–716 (1985).
8. Mahley, R. W. *et al. J. Biol. Chem.* **252**, 7279–7287 (1977).
9. Weisgraber, K. H., Innerarity, T. L. & Mahley, R. W. *J. Biol. Chem.* **253**, 9053–9062 (1978).
10. Koller, E., Quehenberger, O., Jurgens, G., Wolfbeis, O. S. & Esterbauer, H. *FEBS Lett.* **198**, 229–234 (1986).
11. Furchgott, R. F., Jothianandan, D. & Cherry, P. D. in *Vasodilator Mechanisms* (eds Vanhoutte, P. M. & Vatner, S. F.) 1–15 (Karger, Basel, 1984).
12. Brown, M. S. & Goldstein, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3330–3337 (1979).
13. Kannell, W. B., Castelli, W. P., Gordon, T. & McNamara, P. M. *Ann. intern. Med.* **74**, 1–12 (1971).
14. Poole, J. C. F., Sanders, A. G. & Florey, H. W. *J. Path. Bact.* **75**, 133–143 (1958).

Scatter factor is a fibroblast-derived modulator of epithelial cell mobility

Michael Stoker, Ermanno Gherardi,
Marion Perryman & Julia Gray

Department of Pathology, University of Cambridge,
Cambridge CB2 1QP, UK and Imperial Cancer Research Fund,
PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Various factors are known to regulate cell growth and differentiation, but less is known of agents which affect movement and positioning, particularly in epithelial–mesenchymal interactions. Cultured human embryo fibroblasts release a protein with a relative molecular mass (M_r) of ~50,000 (50K) that affects epithelial cells by causing a disruption of junctions, an increase in local motility and a scattering of contiguous sheets of cells. To investigate specificity, a range of cells has been examined for the ability to produce the factor and for sensitivity to its action. Most freshly isolated normal epithelia and epithelia from cell lines of normal tissue, but not epithelia from tumour cell lines or fibroblasts, were sensitive to scatter factor. In contrast, production of the factor, as identified by activity and by chromatography, was restricted to embryonic fibroblasts and certain variants of 3T3 and BHK21 cells and their transformed derivatives. We conclude that the scatter factor is a paracrine effector of epithelial–mesenchymal interaction, which affects the intercellular connections and mobility of normal epithelial cells. The factor might be involved in epithelial migration, such as occurs in embryogenesis or wound healing.

The scattering activity released into culture medium by the MRC5 strain of human fibroblasts was originally observed in human mammary epithelial cells^{1,2}. Subsequently the Madin Darby canine kidney (MDCK) cell line was also found to be highly sensitive, and suitable for assay of activity by observing the pattern of cell dispersal 24 hours after adding the test material to serial dilutions in standard assay plates³ (Fig. 1a, b). Release

of scatter factor from various cells was therefore determined by the standard MDCK cell assay of conditioned medium from confluent cultures. Sensitivity of the cells to scatter factor was examined by time-lapse cinematography to detect rapid dispersal of small colonies of cells, similar to that previously reported with MDCK cells³. In some cases, enhancement of emigration in wounded cultures was also investigated³, particularly when assessing effects on fibroblasts, in which, because they do not form contiguous cell sheets, dispersal was difficult to assess.

Table 1 shows that sensitivity to the scatter factor was restricted to normal epithelial cells, both freshly isolated and as cell lines. Addition of scatter factor from MRC5 cells led, within minutes, to ruffling of the marginal cell membranes and enlargement of colony area. It also led to a reduction in cell density and an increase in the distance between cells, causing cell separation to a degree that varied according to cell type. Of all these sensitive cells, MDCK cells showed the most extensive response, as judged by final cell separation. Nevertheless when the BSC1 monkey line was used to titrate activity, it showed the same sensitivity (titre) as the MDCK cell line, although cell separation was less obvious.

In contrast, no response was seen with a variety of mammary carcinoma lines, normal mammary epithelium transformed by simian virus 40 (SV40) (ref. 4) or lines derived from mouse bladder papillomas. Normal fibroblasts, whether freshly isolated or cell lines and whether producers of factor or not (see below), and vascular endothelial cells, were all unaffected as far as could be seen by change in morphology, movement, or arrangement, or (where tested) by enhancement of migration into wound areas.

Table 1 also shows that scatter activity is only released by fibroblasts and not by the epithelial cells tested, whether they were sensitive to scatter factor or not. These non-yielding epithelial cells included several tumour cell lines and transformed mammary epithelial cells. Thus the factor does not appear to act as an autocrine agent. Several strains of human and mouse embryo fibroblasts and a fibroblastic teratocarcinoma cell line released the activity, but human embryo fibroblasts transformed

Table 1 Production of scatter factor by various cultured cells and sensitivity of cells to its action

Epithelium	Production (titre)	Sensitivity	Fibroblasts	Production (titre)	Sensitivity
Freshly isolated—normal			Freshly isolated—normal		
Human { skin (postnatal)*	<2	+	lung { embryo { (MRC5)	128	-/-
skin (fetal)†	NT	-	(WI38)	64	-/-
ectocervix (adult)‡	NT	+	(DFB1500)	32	-/-
breast (adult)	<2	+	(FL12)**	64	NT
Mouse skin (ref. 10)‡	<2	+	(FS6)**	4	NT
Stable normal lines			Human { skin postnatal†	2	NT
Dog kidney (MDCK)	<2	+/+	breast adult	<2	NT
Monkey kidney (BSC1)	<2	+/+	skin adult (breast cancer)**	<2	NT
Mouse { mammary adult (Comma 1D)§	<2	+/+	Mouse carcass embryo	128	NT
mammary fetal (NMG)§	<2	+/+	Stable normal lines		
bladder (MB 331)	<2	+	Mouse { BALB	<2	-/-
Stable tumour lines			3T3 { Swiss	<2	-
hepatoma HepG2	<2	-/-	Swiss (J2)*	128	-/-
trans SV40 (FR2)¶	<2	-	NIH/1	<2	NT
trans SV40 (FR5)¶	<2	-	NIH/2††	64-256	NT/-
(ZR751)	<2	-	10T1/2	<2	NT
(MDA157)	<2	-	Hamster { BHK21	<2	-
(CAMA 1)	<2	-	Nil8	<2	-
(T47D)	<2	-	Stable tumour lines		
(BT474)	<2	-	Mouse 3T3 NIH/2 clone D4	512	NT/-
(BT20)	<2	-	{ ras††	32	NT
(MCF7)	<2	-	{ myc††	512	NT
melanoma (A2058)*	<2	NT	{ sis††	32	NT
Mouse bladder papilloma { (BB N3)¶	<2	-	{ src††	66	NT
{ (BB N7)¶	<2	-	Hamster BHK 21 trans SV40‡‡	<2	NT
			Rat fibrosarcoma (FS9)§	<2	NT
			Human { teratocarcinoma (Bu)	16	-/-
			lung embryo (MRC5) trans SV40 { (V1)§§	4	NT
			{ (V2)§§	<2	-
			Other cell types		
			Human { vascular endothelium	<4	-
			mesothelium (SP46)¶¶	<2	NT
			Pig chondrocytes*	<4	NT
			Platelet extract**	<2	-

With few exceptions the cells were grown to confluence in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS) (growth medium). To assay production, cells were then exposed for 3 days to DMEM with or without 5% FCS (which had no effect on production), in ~1 ml medium per 10⁶ cells, depending on the size of the dish. The conditioned medium was centrifuged to collect the cells and the supernatant serially diluted in 150 µl volumes of assay medium (DMEM, 5% FCS) in 96-well assay plates, before addition of 3,000 MDCK cells per well in a further 150 µl of assay medium. After overnight incubation, fixation in formaldehyde and Giemsa staining, the degree of scattering was determined and the highest dilution with 30% maximal scattering was recorded as the end point or titre, shown in the 'Production' column. One activity unit as defined as the amount in 300 µl at the end point.

To determine sensitivity, cells were grown to form colonies of 8-64 cells in growth medium. After being moved to assay medium, the colonies were filmed by time-lapse cinematography for at least 6 h before addition of sufficient medium conditioned by MRC5 cells to contain 100 units of scatter factor per ml. Filming of the same field continued for a further 16 h. A positive result, shown as '+' in the right hand column to the left of the slash, indicates a typical increase of membrane ruffling and rapid centrifugal spreading of the colonies, with varying degrees of final cell separation; absence of change is shown as '-'. As an additional test of sensitivity some cells were grown to confluence, wounded³, and changed to assay medium with or without 100 units per ml of scatter factor in conditioned medium. The cultures were examined after 48 h for enhancement of emigration and scattering of cells compared to the controls. Cell lines transformed by SV40 are marked 'trans SV40'. Results are shown as + or - after the slash in the right hand column; NT, not tested. Instead of DMEM, RPMI 1640 medium was used for human mammary epithelium. Some variations may also have occurred when conditioned medium was supplied direct from other laboratories. Cells or conditioned medium were supplied by the following:

* Dr Fiona Watt, Kennedy Institute, London; † Dr Margaret Stanley, Department of Pathology, Cambridge University; ‡ Dr Henry Hennings, NCI Bethesda; § Dr Paul Edwards, Department of Pathology, Cambridge University; || Dr Ian Summerhay, Institute of Cancer Research, London; ¶ Dr Sidney Chang, Royal College Surgeons, London; # Dr Lance Liotta, National Cancer Institute Bethesda; ** Dr Seth Schor, Department of Oncology, Manchester; †† Dr Chris Marshall, Institute Cancer Research London; ‡‡ Dr Henry Rozengurt, Imperial Cancer Research Fund London; §§ Drs Robin Holliday and Lily Hushitscha, NIMR, London; ||| Dr John Gordon, Clinical Research Centre London; ¶¶ Dr Jim Rheinwald, Dana-Farber Cancer Institute, Boston; ** Dr Carl Heldin, Ludwig Institute Uppsala.

by SV40 (ref. 5) yielded reduced or undetectable levels compared to the original cells. Moreover, cultured fibroblasts from post-natal humans also released little or no activity. These included fibroblasts obtained from a patient with cancer, that were reported by Schor *et al.*⁶ to have some characteristics of embryo cells.

Several stock fibroblast cell lines from normal tissues, such as 3T3 and BHK21 cells, released no detectable activity, although some variants of these cells were found to be active producers. Thus BHK21 cells transformed by SV40, unlike human embryo fibroblasts, became modest producers, whereas a clone of Swiss 3T3 cells, selected by Rheinwald and Green⁷ for keratinocyte propagation, gave high levels of scatter factor. Finally, although one line of NIH 3T3 cells (NIH/1) was a non-producer, another line (NIH/2) was an active producer, as were oncogene-transformed derivatives of this line. From this it appears that synthesis or release of the factor from fibroblast lines is variable and under some form of control which is as yet

unidentified. Finally, some other cell types were insensitive and some were non-producers; the latter included the melanoma line A2058 from which Liotta and colleagues have isolated an autocrine motility factor⁸.

The yield per cell from MRC5 cultures is proportional to cell numbers in cultures of different density and although the titres of activity from other types of donor cells varied substantially, this was largely the result of differing cell numbers in the confluent cultures. In 48 hours, MRC5 cells and all the activity-releasing variants of Swiss and NIH 3T3 cells released 400-800 units per 10⁶ cells (the lowest level of detection being 10 units per 10⁶ cells).

A partial characterization of the factor released by the MRC5 strain of human embryo lung fibroblasts, a *ras*-transformed clone (D4) of mouse NIH 3T3 cells and the J2 clone of mouse Swiss 3T3 cells, was obtained by FPLC gel filtration and reverse phase chromatography (RPC). Gel filtration chromatography of serum-free conditioned medium under native conditions

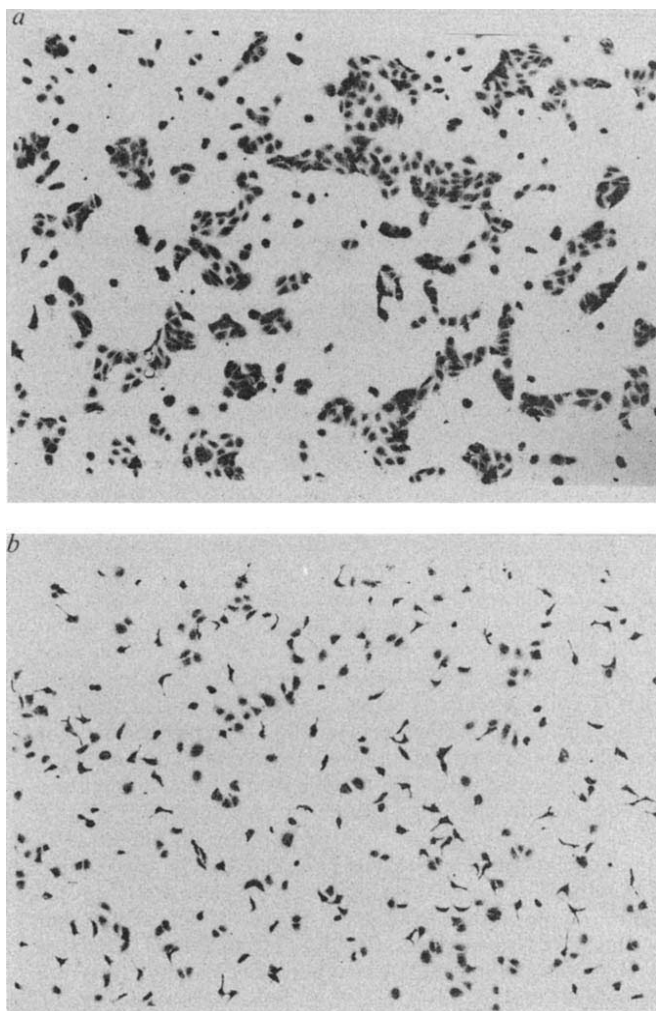
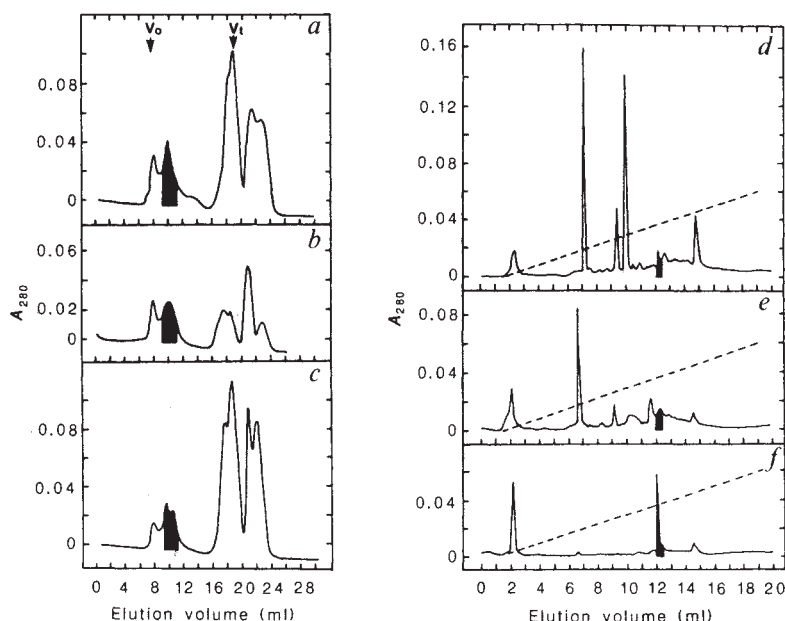


Fig. 1 Assay of scatter factor. *a*, Appearance of MDCK cells in assay plate 16 h after adding 32 units of partially purified fraction from reverse phase FPLC. *b*, Control without addition (Giemsa stain, $\times 40$).

Fig. 2 FPLC gel filtration (*a, b, c*) and reverse phase chromatography (*d, e, f*) of scatter activity released by fibroblasts: *a, d*, MRC5; *b, e*, ras NIH 3T3; *c, f*, Swiss 3T3 J2. Black areas, fractions with activity in the MDCK assay (see Table 1).

Methods. Serum free conditioned medium (50 ml) from each cell type was produced as in Table 1 and concentrated to 2–2.5 ml on Amicon PM30 membranes. Aliquots (0.5 ml) were adjusted to 6 M guanidinium chloride and loaded on a FPLC Superose 12 column (Pharmacia 17-0538-01) equilibrated with 0.05 M Tris-Cl, 6 M guanidinium chloride, pH 8.0. The column was eluted at 0.4 ml min^{-1} and 0.5-ml fractions collected were dialysed for 24 h on a microdialysis unit (BRL M 1200) and their activity measured in the MDCK assay. The column was calibrated with Blue Dextran (V_0), acetone (V_t), galactosidase (114K), bovine serum albumin (67K), egg albumin (45K) and carbonic anhydrase (29K). The peak of scatter activity eluted with an apparent M_r of 50K. For reverse phase chromatography of the scatter activity, 1-ml aliquots of conditioned medium, concentrated as above, were dialysed for 36–48 h against distilled water at 4°C , adjusted with trifluoroacetic acid (TFA) to 1 g l^{-1} and 0.5-ml aliquots loaded on a FPLC ProRPC column (Pharmacia 17-0539-01) equilibrated in TFA, 1 g l^{-1} , in water. The column was eluted with a 0–100% linear gradient of TFA, 1 g l^{-1} , in acetonitrile (Fig. 2*d–f* dotted line) at a flow rate of 0.25 ml min^{-1} . Fractions (0.5 ml) were collected, freeze dried, resuspended in 0.3 ml of phosphate-buffered saline and examined for scatter activity in the MDCK assay and for proteins by SDS-PAGE¹¹.



resulted in low recovery ($<5\%$) of scatter activity, all associated with the void volume peak ($M_r > 500\text{K}$) (data not shown). In the presence of 6 M guanidinium chloride, $\sim 50\%$ of the scatter activity was recovered from a Superose-12 column as a peak with an M_r of about 50K. The elution volume of the activity in conditioned medium from MRC5, ras NIH 3T3, and Swiss 3T3 J2 cells was identical (Fig. 2*a, b* and *c* respectively). Further indication that the scatter activity released by all three cell types is due to the same protein was obtained by RPC on a ProRPC column (Fig. 2*d, e* and *f*) and by anion exchange chromatography on a Mono-Q column (data not shown). Although the factor has not yet been purified, preliminary analysis of the active fractions from all the reverse phase columns by SDS-polyacrylamide gel electrophoresis (PAGE) showed four major bands, ranging in M_r from 40K to 67K.

The factor which we have identified is unusual in several respects. It is cell specific but not species specific: the production and release of the factor is limited to fibroblast cells, but a response has only been observed in epithelial cells. Even allowing for failure to detect low levels of production or alternative responses in our tests, the scatter factor is apparently a paracrine agent and a mediator of fibroblast epithelial interaction. So far the donors are limited to embryo fibroblasts, human lung and mouse carcass, and to a few variants of fibroblast cell lines, but with inconsistent effects of transformation by oncogenes. Whether production is controlled at the level of synthesis or release is not yet known, but the active peptides from three different sources, including human and mouse, are so far indistinguishable by gel filtration, ion exchange, and RPC.

With the exception of the Bu teratocarcinoma cell line, no scatter factor is released from cancer cell lines, either epithelial or fibroblastic, nor are such lines responsive to the factor. Scatter factor also differs in cell specificity from the autocrine motility factor released by melanoma cells⁸ (L. Liotta and M. Strake, personal communication). On the other hand the autocrine motility factor and the scatter factor are both heat-labile proteins of about the same size, and they might both be representatives of a new class of cell-specific endogenous regulators affecting cell migration in both normal and abnormal situations, such as embryogenesis, tumour infiltration, or even wound healing (if production in postnatal fibroblasts can be locally induced). Although there is as yet no evidence that these factors have a function *in vivo*, it is of particular interest that Stern and his

colleagues⁹ have found that when implants of embryo fibroblasts releasing scatter factor were implanted into chick embryos they caused abnormal development whereas non-yielding cells had no effect.

Received 6 February; accepted 12 March 1987.

1. Stoker, M. & Perryman, M. *Molec. Biol. Med.* **1**, 117-122 (1982).
2. Stoker, M. *J. cell Physiol.* **121**, 174-183 (1984).
3. Stoker, M. & Perryman, M. *J. Cell Sci.* **77**, 209-233 (1985).
4. Chang, S. E., Keen, J., Lane, E. & Taylor-Papadimitrou, J. *Cancer Res.* **42**, 2040-2053 (1982).
5. Huschtscha, L. & Holliday, R. *J. Cell Sci.* **63**, 77-99 (1983).
6. Schor, S. L., Schor, A. M., Durning, P. & Rushton, G. *J. Cell Sci.* **73**, 235-244 (1985).
7. Rheinwald, J. G. & Green, H. *Cell* **18**, 331-344 (1975).
8. Liotta, L. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3303-3306 (1986).
9. Ireland, G. K., Stern, C. D. & Stoker, M. *J. Anat.* (in the press).
10. Hennings, H. in *Functional Complexes of Epithelial Cells* (Ciba Foundation Symp. 125) (ed. Bock, G. & Clark, S.) 236-237 (Wiley, Chichester, 1986).
11. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Isolation of an excision product of T-cell receptor α -chain gene rearrangements

Shinji Fujimoto & Hideo Yamagishi

Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

The genes for the T-cell receptor, like the immunoglobulin genes, are rearranged as DNA. The mechanism of this rearrangement is not clear; unequal crossover between chromosomes and the looping-out and excision of the excess DNA have both been suggested. We isolated small polydisperse circular (spc) DNAs from mouse thymocytes and cloned them into a phage vector. Of the 56 clones we analysed, nine contained sequences homologous to T-cell receptor α -chain joining (J_α) segments. We have characterized one of these clones; it contains one J_α segment, and the product out of the recombination of a variable region of the α -chain gene (V_α) with a J_α gene segment. This is the first demonstration of the presence in extrachromosomal DNA of a reciprocal recombination product of any rearranging immunoglobulin or T-cell receptor gene. The finding verifies that V_α - J_α joining can occur by the looping-out and excision of chromosomal DNA.

The presence of extrachromosomal spc-DNA composed entirely of chromosomal sequences seems to reflect the plasticity of eukaryotic genomes^{1,2}. Recently, a population including relatively large spc-DNAs of more than 150 kilobases (kb) was isolated from thymocytes of four-week-old mice (Std: ddY) and cloned into the λ Charon 7 phage or pUC19 plasmid vector³⁻⁶. To examine the possible relationship between the generation of spc-DNAs and the rearrangement of T-cell receptor genes in thymocytes, we screened 56 spc-DNA clones by plaque hybridization using driver DNAs⁷ and appropriate hybridization probes containing sequences derived from the α -chain gene of the T-cell receptor. The insert of the cosmid TA 28.1 (ref. 8) consists of a 45-kb genomic DNA fragment containing a number of J_α segments that include J_1 and J_2 reported previously⁹. This probe hybridized strongly with three clones, MT028, MT113 and MT115, and weakly with six other clones, MT004, MT107, MT112, MT114, MT026 and MT117. An interpretation of these results is that some of the thymus-derived spc-DNA are reciprocal recombination products of the joined V_α - J_α sequences.

We chose clone MT113 for further study and subcloned its insert in a plasmid vector pUC19 because the 7.7-kb insert was the largest found in the J_α probe-positive spc-DNA clones. The restriction map of the 7.7-kb insert is shown in Fig. 1. We wished to locate the putative recombination site on the 7.7-kb DNA; for this purpose the insert was double-digested with restriction

Fig. 1 Restriction map of a p113 spc-DNA insert containing the fragment rearranged from the genomic sequence. Restriction sites are indicated by vertical lines above the horizontal line. Restriction enzymes: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; K, *Kpn*I. The location of a short interspersed repetitive B1 sequence (■) was determined by Southern blot analyses and the locations of J_1 (□) and the 'reciprocal joint' (○) were determined by sequencing the 0.85-kb *Kpn*I-*Hind*III fragment.

enzymes *Hind*III and *Eco*RI. The resulting five DNA fragments were isolated and used as hybridization probes for Southern blots of mouse liver DNA that had been double-digested with the same combination of enzymes (data not shown). In most cases a DNA fragment corresponding in size to the probe used was detected as the major band. One notable exception was found when the 1.2-kb *Eco*RI-*Hind*III fragment occupying the right-most region of the insert depicted in Fig. 1 was used as the hybridization probe. In this case the corresponding 1.2-kb DNA fragment was absent in the Southern blot and a discrete band of 1.45-kb and two faint bands of 1.7 and 1.9 kb were observed. A similar hybridization pattern was obtained when *Eco*RI/*Hind*III double-digested DNA was analysed with the 0.85-kb *Kpn*I-*Hind*III fragment (see Fig. 1) as the hybridization probe. These results strongly suggest that the putative recombination site is located within the 0.85-kb *Kpn*I-*Hind*III fragment. The sequence of this DNA fragment is shown in Fig. 2. A sequence highly homologous to the germline J_1 sequence (a joining segment of the α chain)⁹ was found at nucleotides 24-128. Homology is 96% at the DNA level and 95% at the protein level. Because the various J_α segments identified to date are at most 72% homologous at the protein level¹¹, it is very likely that the J_α segment identified here is J_1 and that the minor nucleotide differences observed are attributable to strain-dependent polymorphism. Immediately 5' to the J_1 -coding sequence are the heptamer and nonamer sequences (indicated in Fig. 2) separated by a 12-base pair (bp) spacer. Sequences highly homologous to these heptamer and nonamer accompany all rearranging gene segments of immunoglobulin and T-cell-receptor genes and have thus been implicated in somatic recombination¹²⁻¹⁴ (heptamer-nonamer pair is called a 'signal sequence'). An additional conserved heptamer and nonamer sequence flanking a 12-bp spacer was found at nucleotides 694-721, possibly derived from the signal sequence that accompanies J_2 , another J_α segment, which has been mapped to ~0.5 kb downstream of J_1 ⁹. We confirmed this by comparing our sequence with the unpublished sequence provided by S. Tonegawa: the two sequences are virtually identical over the entire 0.5-kb region separating the two pairs of heptamer and nonamer, there are some small differences but these are attributable to strain-dependent polymorphism. Although the J_2 -associated signal sequence is found at nucleotide positions 694-721, no J_2 -coding sequence follows it. Instead, a third signal sequence, this time with a 23-bp spacer, is directly attached head-to-head to the J_2 -associated signal sequence. A signal sequence with a spacer of ~23 bp accompanies all V_α gene segments sequenced to date^{8,9}. Indeed the signal sequence and the 23-bp spacer in question are 80% homologous to the 3'-flanking region of one of the V_α gene segments, $V_{\alpha 5H}$ ⁸. Hence we conclude that the sequence spanning the nucleotide positions 694-760 represents two head-to-head joined signal sequences that are generated as a reciprocal recombination product upon the joining of a V_α and J_2 gene segment. Reciprocal joint sequences were also found in MT117, MT028S and MT005 (M. Toda, S.F., K. Tezuka, T. Ohbayashi, T. Iwasato and H.Y., unpublished results).

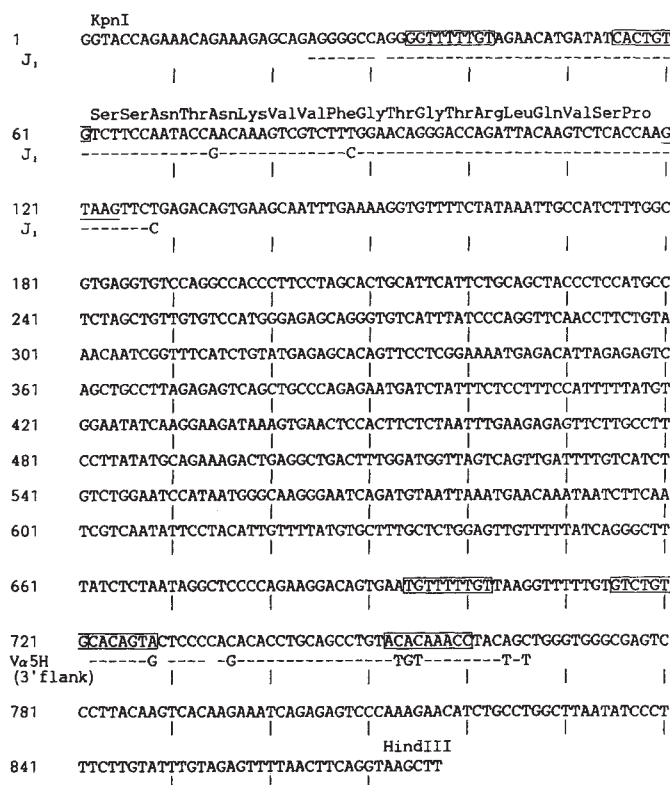


Fig. 2 Nucleotide and protein sequences of the 0.85-kb *KpnI*-*HindIII* fragment of p113 and comparison with the J_1 sequence⁹ and the 3'-flanking sequence of $V_{\alpha 5H}$. The nucleotide sequences were determined by the dideoxy chain-termination method as described in ref. 10. Perfect homology is indicated (for the available sequence data only) by dashes. Signal sequences are boxed and the splicing signal is underlined.

Reciprocal recombination products with a structure similar to that described here have been previously identified for immunoglobulin genes in some myelomas and B cell tumours¹⁵⁻¹⁷ and for a T-cell receptor β -gene in a proliferative T-cell clone¹⁸. But these recombination products were thought to remain as part of chromosomal DNA, implying that unequal cross-over (including a sister-chromatid exchange) or inversion were the probable recombination mechanisms. But evidence has been reported indicating that immunoglobulin gene rearrangement, at least in some cases, involves the deletion of the intervening sequence as a result of a looping-out and excision of this sequence^{19,20}. The discovery of a reciprocal joint on an extra-chromosomal circular DNA molecule supports the looping-out excision model illustrated in Fig. 3. Presumably precise endonucleolytic cutting at the base of the stem and re-ligation generate a circular molecule, sealed at the reciprocal joint, and fuse a V_{α} gene segment to the J_2 segment.

Another interesting feature of spc-DNA in thymocytes is the unusually high copy number (~200 copies per cell⁴) and the enrichment of α -chain gene-related sequences. Out of 56 DNA clones analysed, at least 17 clones showed homology to the V_{α} - or J_{α} -regions (unpublished results). This high copy number of spc-DNA per cell may result from repeated circular DNA formation from multiple recombination events in the V_{α} and J_{α} region or the amplification of spc-DNA by autonomous replication. Final understanding awaits further elucidation of the mechanism and frequency of the α -chain gene rearrangements.

We thank Dr H. Ozeki for his interest, Drs T. Honjo and K.

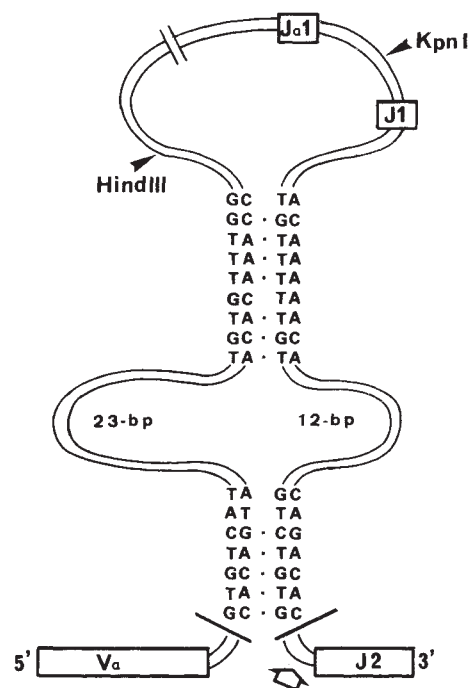


Fig. 3 Inverted repeat stem structure formed between 3'-non-coding regions of embryonic V_{α} DNA and the 5'-flanking region of α -chain J_2 DNA. Coding regions are boxed. $J_{\alpha 1}$ is located in the most 5' region of the J_{α} gene cluster^{8,21}. The cleavage activity (oblique lines) and binding activity (dots) of a putative recombinase and the activity for the 'junctional' diversification of a coding joint (open arrow) are shown. Both ends of the 0.85-kb *KpnI*-*HindIII* fragment of the spc-DNA clone p113 are marked by arrow heads.

Ikuta for discussions. Dr L. Hood for the T-cell receptor gene probes and Dr S. Tonegawa for unpublished sequence data and advice. We also thank M. Toda, K. Tezuka, T. Ohbayashi and T. Iwasato for performing the plaque hybridization experiments. This work was supported by Grants-in-Aid for Science Research and Cancer Research from the Ministry of Education, Science and Culture of Japan, and by a Shimizu Foundation Research Grant.

Received 20 December 1986; revised 1 April; accepted 22 April 1987.

1. Rush, M. G. & Misra, R. *Plasmid* **14**, 177-191 (1985).
2. Yamagishi, H. *BioEssays* **4**, 218-221 (1986).
3. Yamagishi, H., Kunisada, T. & Tsuda, T. *Plasmid* **8**, 299-306 (1982).
4. Tsuda, T. *et al. Plasmid* **10**, 235-241 (1983).
5. Yamagishi, H. *et al. Gene* **26**, 317-321 (1983).
6. Fujimoto, S., Tsuda, T., Toda, M. & Yamagishi, H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2072-2076 (1985).
7. Sealey, P. G., Whittaker, P. A. & Southern, E. M. *Nucleic Acids Res.* **13**, 1905-1922 (1985).
8. Winoto, A., Mjolsness, S. & Hood, L. *Nature* **316**, 832-836 (1985).
9. Hayday, A. C. *et al. Nature* **316**, 828-832 (1985).
10. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
11. Arden, B., Klotz, J. L., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).
12. Tonegawa, S. *Nature* **302**, 575-581 (1983).
13. Honjo, T. *A. Rev. Immun.* **1**, 499-528 (1983).
14. Malissen, M. *et al. Cell* **37**, 1101-1110 (1984).
15. Hochtl, J., Muller, C. R. & Zachau, H. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1383-1387 (1982).
16. Alt, F. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
17. Deev, S. M., Combrato, G., Klobbeck, H.-G. & Zachau, H. G. *Nucleic Acids Res.* **15**, 1-14 (1987).
18. Malissen, M. *et al. Nature* **319**, 28-33 (1986).
19. Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
20. Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6022-6026 (1980).
21. Ikuta, K. *et al. J. exp. Med.* **164**, 428-442 (1986).

Antibodies against the paramyxovirus SV5 in the cerebrospinal fluids of some multiple sclerosis patients

K. K. A. Goswami*, R. E. Randall*, L. S. Lange† & W. C. Russell*‡

* Department of Biochemistry and Microbiology, University of St Andrews, St Andrews, Fife KY16 9AL, UK

† Department of Neurology, Charing Cross Hospital, London W6 8RF, UK

Multiple sclerosis is a demyelinating disease of the central nervous system and although there is little doubt that an infectious agent (or agents) is involved (for reviews see refs 1 and 2) it has not been possible to demonstrate this unequivocally by any direct relationship to a given agent (see, for example refs 3–5). Here we show that a significant proportion of patients with multiple sclerosis have antibodies against the paramyxovirus SV5 (simian virus 5) in their cerebrospinal fluid and in some of these such antibodies form a major proportion of the total immunoglobulin content. Further, we have been able to demonstrate that the oligoclonal bands displayed on electrophoresis of the cerebrospinal fluid of these patients can be removed by prior absorption with SV5 virus antigen.

The paramyxovirus SV5 was originally detected in monkey tissues⁶ and has never been associated with any human disease. It has been isolated from a number of human sources^{7,8}, however, including the bone marrows of multiple sclerosis (MS) patients^{9,10}. Subsequent studies, using monoclonal antibodies¹¹, showed that the virus could be detected directly in the bone marrows of both MS and non-MS patients as a persistent infection¹². Serological studies also demonstrated that about 40% of individuals tested had been infected with SV5 although no major differences were detected in the prevalence or levels of SV5 antibodies in the sera of MS compared to non-MS patients¹³.

We have now extended these analyses to the cerebrospinal fluids (CSFs) of a series of MS and non-MS patients and have found a much clearer correlation between the presence of specific SV5 antibodies in the CSFs and the disease. The MS patients were assessed using the internationally accepted criteria for diagnosis. The controls were patients with motor neuron disease, cerebrovascular disease, glioma, neurosarcooidosis, polycythemia, migraine, cerebral haemorrhage and other neurological disorders. Only a limited number of assays could be attempted on each CSF because of the small volumes (100–500 µl) available in many cases. Tests were carried out before the diagnoses of the patients were revealed to the experimenters. In the first series of tests, the CSFs of 34 MS and 30 non-MS patients were analysed for antibodies against SV5, the closely related human parainfluenza virus type 2 (PF-2), measles virus and adenovirus. The initial screening for virus antibodies was by immunofluorescence¹¹ followed at a later stage by radioimmune precipitation¹⁰. Because of cross-reacting antigens, immunofluorescence does not differentiate SV5 and PF-2 antibodies. However they can be distinguished by specific antigenic determinants on the major surface glycoprotein, the haemagglutinin-neuraminidase (HN) polypeptide. Consequently, where appropriate, radioimmune precipitation was used^{10,11} to identify the antibodies against the HN polypeptides. A significant proportion of the CSFs of MS patients contained SV5-specific antibodies (19 out of 34; detected by immunofluorescence, confirmed by radioimmune precipitation) but a small number of control CSFs (7 out of 30 by immunofluorescence but only 3 of these were positive by radioimmune precipitation) also possessed such antibodies. The three control patients who

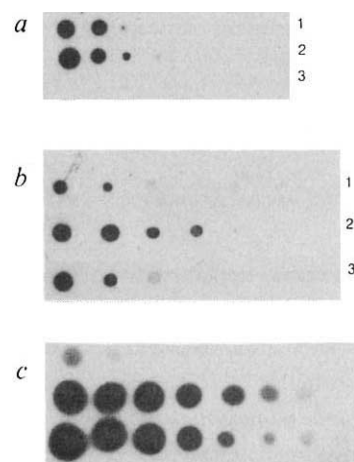


Fig. 1 Assay for SV5 immunoglobulin in CSFs. *a*, CSF from MS patient (SV5 immunofluorescence (FA) titre 1/5,000). Purified adenovirus type 5 (ref. 25) (Ad5, 1×10^{10} PFU ml⁻¹) was added to one aliquot of CSF as an additional control. SV5 virus apparently removed all detectable IgG. Row 1, CSF plus PBS; row 2, CSF plus Ad5; row 3, CSF plus SV5. *b*, CSF from MS patient (SV5 FA titre 1/500). A standard preparation of IgG (first dilution 6 mg dl⁻¹) was also titrated in parallel. Row 1, CSF plus SV5; row 2, CSF plus PBS; row 3, standard IgG. *c*, CSF from MS patient (SV5 FA titre 1/100). In this case antibody specific for SV5 was measured by soaking the nitrocellulose paper in 1 ml of purified SV5 virus in PBS and washing in PBS containing 3% bovine serum albumin (BSA) for 20 min at 37 °C before blotting dry and then adding dilutions as described below. Purified adenovirus type 5 was used as an additional control. Row 1, CSF plus SV5; row 2, CSF plus Ad5; row 3, CSF plus PBS.

Methods. Aliquots (300 µl) of CSF samples were clarified by centrifugation at 10,000g for 20 min at 4 °C (MSE MicroCentaur). The supernatant was removed and divided into aliquots of 50 µl. Equal volumes of purified SV5 virus¹⁰ or PBS were added to each aliquot in a micro-Eppendorf tube, mixed well and then left overnight at 4 °C. The mixtures were then centrifuged (10,000g for 1 h) at 4 °C and 50 µl of each supernatant was carefully withdrawn by a capillary pipette and doubling dilutions were made in PBS in a microtitre plate. Aliquots (3 µl) of each dilution were then added to nitrocellulose paper BA85 (Schleicher and Schell) marked in 1-cm squares and incubated in a flat plastic box for 20 min at 37 °C in PBS containing 3% BSA. The paper was then washed gently with PBS containing 0.2% Nonidet P40 (NP40) followed by immersion in 5 ml of PBS containing ¹²⁵I-labelled protein A¹¹ (10⁶ c.p.m.) for 1 h. After being washed twice for 15 min with PBS, the nitrocellulose sheet was dried and mounted for autoradiography.

were positive were suffering from neurological conditions where blood-brain barrier (BBB) damage was likely. When the CSFs and sera were analysed for albumin content it was evident that these three CSFs were those showing the most severe BBB disruption.

In general agreement with previous findings^{14,15}, we detected antibodies against both measles virus and PF-2 in MS and control CSFs. Significantly, although high levels of antibodies against adenovirus could be detected in sera, only the three neurological control patients described above, showing severe BBB damage, gave evidence of adenovirus antibodies in CSF. Thus we conclude that the presence of SV5 antibodies in the CSFs of patients in the absence of severe BBB damage correlates well with MS. No such close disease relationship was noted with other virus antibodies in our tests (patient and serological data will be published elsewhere). These observations alone do not necessarily imply a role for SV5 in the aetiology of MS; we explored this aspect further by examining the nature and amount of the SV5 antibody in the CSF.

The nature of the SV5 antibodies in the CSFs was most readily analysed by immunoprecipitation of extracts of radiolabelled

‡ To whom correspondence should be addressed.

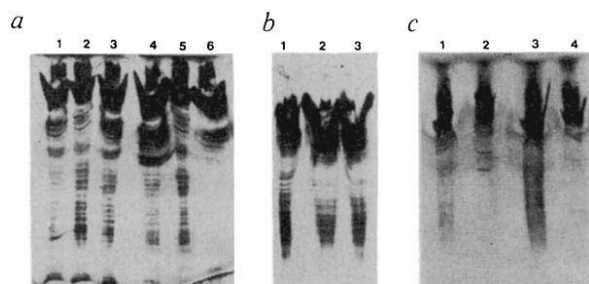


Fig. 2 Analysis of the oligoclonal bands of electropherograms after isoelectric focusing and silver staining from the CSFs following absorption with SV5 and measles virus preparations. *a*, Lane 1, CSF from MS patient with negative SV5 antibody and PBS; lane 2, CSF from the same MS patient plus measles virus; lane 3, CSF from the same MS patient plus SV5; lane 4, CSF from MS patient (SV5 FA titre 1/100) plus PBS; lane 5, CSF from the same MS patient plus measles virus; lane 6, CSF from the same MS patient plus SV5. *b*, Lane 1, CSF from control (neurosarcoidosis) patient (SV5 FA titre 1/50) plus SV5; lane 2, CSF from the same patient plus purified PF-3 virus¹¹; Lane 3, CSF from the same patient plus PBS. *c*, Lane 1, CSF from MS patient with positive SV5 antibody (SV FA titre 1/100) plus PBS; lane 2, CSF from the same MS patient plus SV5 virus; lane 3, CSF from (a different) MS patient with positive SV5 antibody (SV5 FA titre 1/100) plus PBS; lane 4, CSF from the same MS patient plus SV5. PF-3 virus failed to absorb out the oligoclonal bands from these two patients (data not shown). The SV5 and PF-3 viruses were purified by density gradient centrifugation^{10,11}. The pH gradients measured by electrophoresis of IEF standard proteins in parallel showed pH of 4.6 at top to pH 9.0 at bottom. The major band at the top is albumin.

Methods. Measles (Phil/26 strain) and SV5 (LN strain^{9,13}) viruses were prepared by infecting (at a multiplicity of 0.1 PFU per cell) vero cells (a line of African green monkey kidney cells) growing as monolayers at 37°C in rotating winchester bottles. All virus seeds had been plaque purified and were mycoplasma free—as were the vero cells (using both culture and staining procedures²⁶). When cytopathic effect was just evident the culture medium was replaced with medium deprived of calf serum and incubation continued for a further 8 h. The medium was then collected and after low-speed centrifugation (5,000g for 10 min) to remove cellular debris the virus was pelleted by high-speed centrifugation (100,000g for 1 h) and the pellet resuspended in phosphate buffered saline (PBS). Polypeptide profiles of these pellets by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining showed that the suspensions consisted of ~20–60% virus proteins (for example, SV5 in Fig. 3, lane 1). For the absorption experiments, the virus suspensions were each adjusted to give a protein content of ~5 mg ml⁻¹. Before absorption, 300 µl aliquots of the CSFs were centrifuged at 350,000g for 2 h in 1-ml polypropylene tubes in the TL-100 ultracentrifuge (Beckman) at 4°C. Aliquots (95 µl) of the supernatants were then dispensed into small Eppendorf tubes and 5-µl aliquots of the SV5 or measles antigens (prepared as described) or PBS were added to each of the tubes and the contents of the three tubes well mixed and left overnight at 4°C to absorb. The CSF-antigen mixtures were then transferred to 1-ml centrifuge tubes and centrifuged again in the ultracentrifuge at 350,000g for 2 h. About 70 µl of each supernatants was then carefully removed with capillaries and aliquots were submitted to analytical isoelectric focusing using the Multiphor II (LKB) on ampholine PAG plates (pH 3.5–9.5). The samples were electrophoresed for a total of 105 min under standard conditions with a constant voltage (1,000 V) and an initial current of 30 mA. The PAG plates were then subjected to silver staining (Bio Rad Kit) and photographed.

SV5-infected cells. Most of the CSFs analysed contained antibodies against the HN and the nucleoprotein (NP) polypeptides of the virus although at least one CSF also precipitated other components, possibly matrix and fusion polypeptides¹¹. In the case of one patient (D3) the CSF precipitated only the SV5 HN polypeptide and this is discussed further below (see Fig. 3).

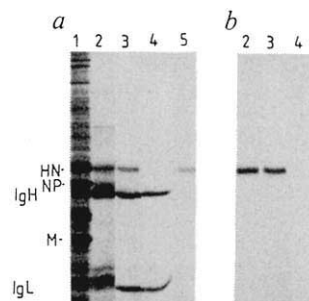


Fig. 3 Electrophoretic analysis by SDS-PAGE of SV5 polypeptides and immunoprecipitates. Electropherograms of *a*, total polypeptides (Coomassie Brilliant Blue stained gel) and *b*, ³⁵S-methionine-labelled polypeptides (autoradiogram) after electrophoresis on a 10% SDS polyacrylamide gel. Lane 1, partially purified preparation of SV5 virus (see Fig. 2 legend) lanes 2–4, immunoprecipitates¹¹ formed by the reaction of either the CSF of an MS patient (D3) (lane 2) or monoclonal antibody to the HN protein of SV5 (lane 3) or a control (non-SV5) monoclonal antibody (lane 4) with a labelled soluble antigen extract from SV5 infected cells; lane 5, a sample of a purified preparation (prepared as in Fig. 4 legend) of the HN polypeptide. HN, NP and M (matrix) indicate the corresponding SV5 virus polypeptides. IgH and IgL indicate immunoglobulin heavy and light chains.

A few of the MS CSFs elaborated quite substantial titres of SV5 antibodies (1/500–1/10,000) and we showed by a radio-immune assay that a significant proportion of the total CSF immunoglobulin (binding to protein A) could be ascribed to the virus antibodies in some of these cases. Figure 1*a* shows the results of an absorption assay where aliquots of one such CSF were incubated with partially purified SV5 virus and with purified adenovirus type 5 virus as a control. Following pelleting of immune complexes, the supernatants were tested for relative immunoglobulin G (IgG) content using ¹²⁵I-labelled protein A as a probe. Figure 1*b* shows the result obtained with another CSF. In this experiment a control IgG was analysed in parallel and the relative intensities of the diluted spots on the autoradiographs suggests that the CSF contained about 12 mg dl⁻¹ of IgG of which at least half could be removed by preabsorption with SV5 virus. Figure 1*c* illustrates the efficacy of the absorption and centrifugation technique in removing antibody specific to SV5 virus from another MS CSF. In this case about 95% of the SV5-specific antibody was removed by the absorption procedures.

For the next stage of the investigation we submitted appropriate CSF samples to electrophoretic analysis and examined the effect on the oligoclonal banding patterns of preabsorption with virus preparations followed by high-speed centrifugation. The CSFs analysed were derived from the residues of material from the initial antibody studies on patients from the London area supplemented by fresh CSFs from patients in Dundee. We used isoelectric focusing of unconcentrated CSFs on thin polyacrylamide gels. Normally an oligoclonal banding pattern could be readily visualized following silver staining of about 2–5 µg of CSF immunoglobulin (in 20 µl)¹⁶. The effect of preabsorption with partially purified preparations on the oligoclonal patterns of the CSFs was then studied. Preparations of measles virus and in some cases PF-3 virus were used as control antigens.

Figure 2*a* shows the results of such an experiment with the CSFs from two MS patients, one with no detectable SV5 antibody in the CSF and the other with SV5-specific antibody. Lanes 1, 2 and 3 show the identical patterns obtained with the CSF without SV5 antibody (lane 1 is a PBS control) after absorption with measles virus (lane 2) and SV5 virus (lane 3). However SV5 virus removed most of the oligoclonal bands from the CSF containing SV5 antibody (lane 6), measles virus (lane 5) having no such effect. Figure 2*b* shows the result of a similar experiment using the CSF from a neurological control (neurosarcoidosis)



Fig. 4 Oligoclonal banding patterns of CSF from patient D3 (SV5 FA titre 1/100) analysed by isoelectric focusing and silver staining after absorption with SV5, measles and a purified preparation of the SV5 HN polypeptide. Lane 1, CSF plus SV5 virus; lane 2, CSF plus measles virus; lane 3, CSF plus purified SV5 HN polypeptide; lane 4, CSF + PBS. The HN polypeptide was purified by immunoaffinity chromatography in which a monoclonal antibody to the SV5 HN polypeptide (R.E.R., D. Young, K.K.A.G. and W.C.R., manuscript in preparation) was linked to CNBr-activated Sepharose-4B by standard procedures. Vero cells were grown and infected with SV5 (as described in Fig. 2 legend for purification of SV5 virus) and when the cells showed 90% cytopathic effect they were disrupted with an ultrasonic probe in lysis buffer (0.5% NP40 in PBS made to 0.65 M NaCl; 5 ml per 2×10^8 cells). A soluble antigen extract was prepared by pelleting the particulate material from the lysed cells by centrifugation at 400,000g in the Beckman TL100 centrifuge. This soluble antigen extract (5 ml) was passed through a 2-ml immunoaffinity column and, after the column had been extensively washed in immune lysis buffer, the bound HN polypeptide was removed from the column with a low-pH buffer (0.1 M glycine HCl pH 2.8) including 0.1% NP40. The eluted preparation of HN polypeptide was then neutralized with 1 M Tris/HCl to pH 7 and checked for purity by SDS-PAGE (see Fig. 3). In the absorption experiment, 5 μ l ($\sim 1 \mu$ g) of the preparation was added to 95 μ l of CSF and processed with the other samples as described in Fig. 2 legend.

patient with significant levels of SV5 antibody. This CSF clearly gives rise to oligoclonal bands but in this case the bands were not removed by absorption with either measles virus or SV5. This may be because the SV5 antibody detected was derived by leakage from the peripheral system and was presumably polyclonal in nature because this patient showed severe BBB damage as judged by albumin analysis (see above). Figure 2c shows similar positive results obtained with two other CSFs in a different series of experiments which used human parainfluenza type 3 as control antigen.

We have analysed 14 CSFs from 12 MS and 2 non-MS patients by this technique and we have been able to show specific absorption of the oligoclonal bands of 8 out of 12 MS CSFs with the SV5 antigen. All these 8 CSFs showed the presence of SV5 antibody as judged by radioimmune precipitation. We were unable to absorb out the oligoclonal bands from the CSF of another MS patient which contained antibodies to SV5/PF-2 (immunofluorescence titre of 1/10), and further studies were not feasible owing to lack of sample. Three CSFs lacking antibodies against SV5 from MS patients did not change their banding patterns after absorption. It is interesting that one of these, although negative for the SV5-specific HN polypeptide, did immunoprecipitate the NP polypeptide. Two CSFs from non-MS patients (neurosarcoidosis and glioma), which were positive for SV5 antibody and showed oligoclonal banding patterns, did not show any changes after virus absorption.

As noted above, the radioimmune precipitation assay on the CSF from one patient (D3) had indicated that apparently only antibody against the SV5 HN polypeptide was present. Figure 3 shows the total protein stain of a polyacrylamide gel and the corresponding autoradiogram from electrophoresis of immune precipitates obtained following the reaction of a soluble antigen extract of SV5-infected cells with the CSF from the patient D3,

in comparison to the immunoprecipitate obtained using a monoclonal antibody against the SV5 HN polypeptide. The protein staining also revealed the immunoglobulin heavy and light chains of both the CSF and the monoclonal antibodies as well as the precipitated HN polypeptide. The relative stoichiometries of the stained bands and the intensities of the autoradiograph bands are consistent with our initial observations (Fig. 1) that the CSF immunoglobulins in some cases appear to be mainly against the SV5 component. Figure 4 shows that the oligoclonal bands could be removed with the purified SV5 HN polypeptide as well as with the SV5 virus preparation. We propose to carry out similar absorption experiments using HN and other purified SV5 polypeptides on other CSF samples when they become available.

A number of other neurological diseases with clear relationship to virus infections also show CSF oligoclonal banding patterns which can be removed by prior absorption of the CSFs with the appropriate virus antigen. Thus in herpes and mump encephalitis, and in subacute sclerosing panencephalitis, bands could be removed by prior absorption with herpes, mumps and measles antigens respectively¹⁷⁻¹⁹. Previous attempts to remove the oligoclonal bands appearing in the CSFs of MS patients by absorption with a variety of viruses (particularly by measles virus¹⁹) have proved unsuccessful. The use of very sensitive immunoassay systems has shown that antibodies against most of the common human viruses could be detected in the CSFs of MS patients. It has therefore been postulated¹⁴ that most of these antibodies have arisen nonspecifically in MS as a consequence of a primary malfunction of the central nervous system (CNS) immune system and many workers think that they represent an epiphenomenon. In contrast, our studies suggest that the paramyxovirus SV5 may be important in the pathogenesis of at least some cases of MS.

The significance of the fact that some MS patients did not appear to have SV5 antibodies in their CSFs remains to be established. Inappropriate sample handling in some cases may provide an explanation, as we have little knowledge of the detailed history of many of the samples. On the other hand, it is clear that the role of other viruses—particularly members of the Paramyxoviridae (for a review see ref. 20)—should be re-examined by the approaches used in this study. Note that the SV5 virus strains used here (RQ and LN) were originally isolated from the bone marrow of MS patients^{9,10,12} and no attempts have been made to absorb out oligoclonal bands with the prototype SV5 virus (or indeed with PF-2 virus). Previous studies have also demonstrated that the SV5 specific HN-1 monoclonal antibody can cross-react in fluorescence and other tests with a component in rat and human brain²¹. These results and others, particularly those showing that the SV5 HN antibody can itself be the major component of the CSF oligoclonal bands, therefore also suggest that the demyelination which appears to be the main lesion in MS might be an autoimmune reaction triggered by the SV5 HN antibody prevalent in the CSF. However, previous efforts to remove immunoglobulin in the CSF by absorption with CNS material have not been successful²².

Because SV5 virus has never been considered to be of any significance in human infections, very little is known of its prevalence in human populations. As indicated above, our own limited serological studies on individuals, mostly in the London area, suggest that $\sim 40\%$ of individuals have been infected with the virus^{12,13}. It will clearly be of some interest to compare prevalences in populations exhibiting high and low incidences of MS. The possibility of a zoonotic infection²³ also cannot be dismissed because SV5 and SV5-like viruses are reported to infect a variety of species⁷. Thus SV5 virus is known to cause a respiratory canine infection²⁴ and a number of canine vaccines incorporating the virus are in use.

The UK Multiple Sclerosis Society supported K.K.A.G. through a grant to W.C.R. and L.S.L. R.E.R. is the recipient of a University Award from the Wellcome Trust (to W.C.R.). We

thank Dr Duncan Davidson, Neurology Department, Dundee Royal Infirmary and Dr E. J. Thompson, National Hospital, London, for clinical material; Dr D. N. Mitchell for support and encouragement; Dan Young for technical assistance; W. Blyth for photography; and M. Wilson for typing.

Received 19 January; accepted 18 March 1987.

1. Cook, S. D. & Dowling, P. C. *Neurology* **30**, 80-91 (1980).
2. Ter Meulen, V. & Stephenson, J. R. in *Multiple Sclerosis* (eds Hallpike, J. F., Adams, C. W. M. & Tourtellotte, W.) 241-274 (Chapman and Hall, London, 1983).
3. Kaprowski, H. *et al. Nature* **318**, 154-160 (1985).
4. Hauser, S. L. *et al. Nature* **322**, 176-177 (1986).
5. Karpas, A., Kämpf, U., Sidén, A., Koch, M. & Poser, S. *Nature* **322**, 177-178 (1986).
6. Hull, R. N., Minner, I. R. & Smith, J. W. *Am. J. Hyg.* **63**, 204-215 (1956).
7. Hsuing, G. D. *Prog. med. Virol.* **14**, 241-247 (1972).
8. Robbins, S. J., Wrzowski, H., Kline, A. L., Tenser, R. B. & Rapp, F. J. *infect. Dis.* **143**, 396-402 (1981).
9. Mitchell, D. N. *et al. Lancet* **ii**, 387-391 (1978).
10. Goswami, K. K. A. & Russell, W. C. *J. gen. Virol.* **60**, 177-183 (1982).
11. Goswami, K. K. A. & Russell, W. C. *J. gen. Virol.* **64**, 1663-1672 (1983).
12. Goswami, K. K. A., Lange, L. S., Mitchell, D. N., Cameron, K. R. & Russell, W. C. *J. gen. Virol.* **65**, 1295-1303 (1984).
13. Goswami, K. K. A., Cameron, K. R., Russell, W. C., Lange, L. S. & Mitchell, D. N. *J. gen. Virol.* **65**, 1881-1888 (1984).
14. Salmi, A., Reunanen, M., Ilonen, J. & Panelius, M. *Clin. exp. Immun.* **52**, 241-249 (1983).
15. Norrby, E. *et al. Infect. Immun.* **10**, 688-694 (1974).
16. Strangaitis, S. M., Shapshak, P., Tourtellotte, W. W., Lee, M. M. & Reiber, H. O. *Electrophoresis* **6**, 287-291 (1985).
17. Vandvik, B. *et al. J. Neurol.* **231**, 307-312 (1985).
18. Nordal, H. J. *et al. Scand. J. Immun.* **7**, 381-386 (1978).
19. Vandvik, B., Norrby, E., Nordal, H. J. & Degro, M. *Scand. J. Immun.* **5**, 979-992 (1976).
20. Pringle, C. R. in *Molecular Basis of Virus Disease. S.G.M. Symposium 40* (eds Russell, W. C. & Almond, J.). (Cambridge University Press, in the press).
21. Goswami, K. K. A., Morris, R. J., Rastogi, S. C., Lange, L. S. & Russell, W. C. *J. Neuroimmun.* **9**, 99-108 (1985).
22. Chow, C.-H. J. *et al. Neurology* **33**, 24-28 (1983).
23. Russell, W. C. in *Immunology of Nervous System Infections, Progress in Brain Research 59* (eds Behan, P. O., ter Meulen, V. & Clifford Rose, F.) (Elsevier, London, 1983).
24. Cornwell, H. I. C., McCandish, I. A. P., Thompson, H., Laird, H. M. & Wright, N. G. *Vet. Rec.* **98**, 301-303 (1976).
25. Russell, W. C., Patel, G., Precious, B., Sharp, I. & Gardner, P. S. *J. gen. Virol.* **56**, 393-408 (1981).
26. Russell, W. C., Newman, C. & Williamson, D. H. *Nature* **253**, 461-462 (1975).

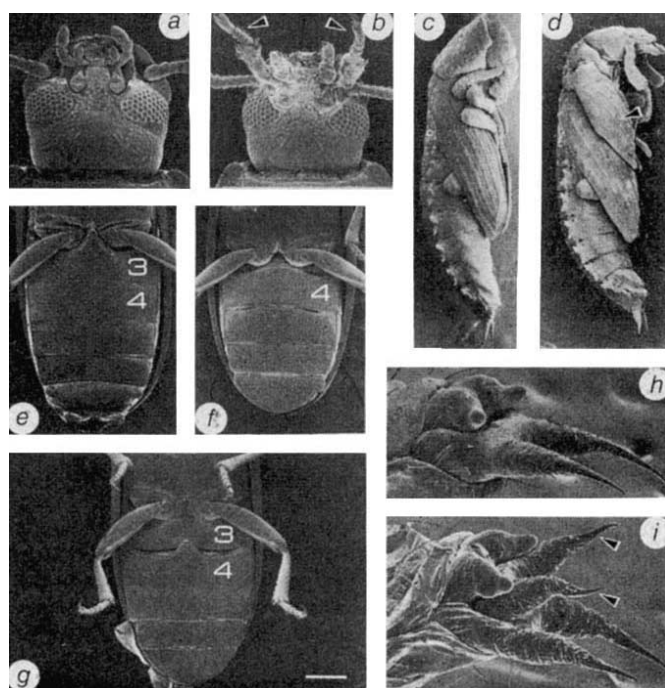


Fig. 1 Mutant phenotypes representing five closely linked homoeotic loci. *a*, Head of wild-type adult, ventral view, showing maxillary (outer) and labial (inner) palps. *b*, Same view of *maxilopodia* (*mxp/mxp^{NG}*). Arrows, transformed maxillary palps. *c*, Wild-type pupa, side view, showing pronotum (hood-like structure behind the head) and elytra. *d*, Same view of *alate prothorax* (*apt/Cx⁶*). Arrow, reiterated elytra extending from transformed pronotum. *e*, Abdomen of wild-type adult, ventral view. The third and fourth sternites are shown by numbers. (Sternites 1 and 2 are modified to form the sockets for the metathoracic coxae, and are not visible.) *f*, Same view of *missing abdominal sternites* (*mas/mas*), revealing the absence of sternites 1-3. *g*, Same view of *pointed abdominal sternite* (*pas³⁰/pas*), revealing the transformation of the fourth sternite into a copy of the third. *h*, Posterior extremities of wild-type female pupa, three-quarter ventral view, showing female papillae (blunt-tipped) and urogomphi (pointed). *i*, Same view of *extra urogomphi* (*eu/eu*) female. Arrows, supernumerary urogomphi. The scale bar in *g* indicates 0.2 mm in *a* and *b*, 0.5 mm in *c* and *d*, 0.33 mm in *e-g*, and 0.1 mm in *h* and *i*. Mutant alleles *notched genae* (*mxp^{NG}*), *pas³⁰*, and *Cephalothorax* (*Cx⁶*) are recently discovered alleles of *mxp*, *pas* and *apt*, respectively.

A homoeotic gene cluster in the red flour beetle

Richard W. Beeman

US Grain Marketing Research Laboratory, US Department of Agriculture, 1515 College Avenue, Manhattan, Kansas 66502, USA

The bodies of insects consist of a tandem array of segments. Homoeotic genes direct cells in different segments to the appropriate developmental pathways. I have discovered the existence of a cluster of homoeotic genes in the flour beetle *Tribolium castaneum* that are ordered along the second linkage group in a sequence identical to that of the body segments in which they act. The known domain of function of this homoeotic gene cluster (HOM-C) spans approximately 15 segments in the head, thorax and abdomen. The HOM-C contains at least six complementation groups, including elements with apparent homology to the two major groups of homoeotic genes in *Drosophila melanogaster*, the Antennapedia complex (ANT-C) and the bithorax complex (BX-C).

In the fruit fly *Drosophila melanogaster*, homoeotic genes are organized into two major clusters, the Antennapedia complex (ANT-C)¹ and the bithorax complex (BX-C)²⁻⁶, both located on the right arm of the third chromosome. The problem of how these and other genes interact to orchestrate normal segment differentiation in *Drosophila* is being examined in many laboratories by a combination of developmental, genetic and molecular approaches¹⁻¹⁰.

The generality of principles learned in *D. melanogaster* is uncertain because little is known about homoeotic genes in other species. Homoeotic mutations have been described sporadically in other insects including *Musca*, *Aedes*, *Anopheles*, *Blattella* and *Tribolium*^{11,12}. A tightly linked series of homoeotic muta-

tions affecting thoracic and abdominal segments has been described in the silkworm *Bombyx mori*¹³.

Mature *Drosophila* embryos are simple, relatively undifferentiated vermiform larvae with few cuticular landmarks that might help to define homoeotic transformations. Our understanding of the genetics of development in segmented animals would be greatly enhanced by detailed analysis of phylogenetically more primitive species, which typically have a prolonged embryonic stage, giving rise to a larva that is more highly differentiated, particularly in the head and thorax, but that nevertheless better retains the integrity and identity of its anterior pattern of segmentation established early in embryogenesis. Tenebrionid flour beetles satisfy these conditions.

I have initiated an investigation of the organization and function of homoeotic genes in the red flour beetle *Tribolium castaneum* (Tenebrionidae). A number of spontaneous mutations have been described and assigned to linkage groups in this species^{14,15}. We have recently shown that visible mutations and chromosomal rearrangements can be readily isolated in *Tribolium* using ionizing radiation and a chemical mutagen (ref. 16 and data not shown). At least 10 homoeotic mutations have been reported in *T. castaneum*^{14,15,17}; their properties are summarized in Table 1. They represent at least seven loci including

Table 1 Homoeotic mutants of *Tribolium castaneum*

Mutant	Origin (ref.)	Linkage group	Properties	Complementation group	Transformed segment(s)	Nature of transformation	Stage affected	Possible <i>Drosophila</i> analogue
<i>ap</i>	S (20)	8	r, HVF	<i>ap</i>	antennal	antenna → leg	L, P, A	<i>ss</i> ^a
<i>mxp</i>	S (21)	2	r, HVF	<i>mxp</i>	maxilla, labium	max. + lab. palp → leg	L, P, A	<i>pb</i>
<i>Dch</i>	X (22)	2	D, HVS	<i>mxp</i>	maxilla	max. palp → leg	P, A	?
(<i>lp</i>)	S (23)	?	?	?	labium	lab. palp → leg	L, P, A	<i>pb</i>
<i>ptl</i>	S (24)	2	r, HVF	<i>ptl</i>	T1	suppresses development	L, P, A	?
<i>apt</i>	S (25)	2	r, HVF	<i>apt</i>	T1	T1 → T2	L, P, A	<i>Scr</i>
<i>mas</i>	S (26)	2	r, HVF	<i>mas</i>	A1-A3	suppresses development	A	?
<i>pas</i>	S (27)	2	r, HVF	<i>pas</i>	A4-A6	A4, 5 → A3 A6 → A7	A	<i>iab-4</i>
<i>eu</i>	S (17)	2	r, HVF	<i>eu</i>	A10	A10 → A9	L, P	?
(<i>Tu</i>)	S (28)	?	D, HL	?	A10	A10 → A9	L, P, A	?

Names of mutants are as described below; Other abbreviations: *ap*, *antennapedia*; S, spontaneous; X, X-ray-induced; r, recessive; D, dominant; HVF, homozygous viable and fertile; HVS, homozygous viable but sterile; HL, homozygous lethal; T1, prothorax; T2, mesothorax; A1-10, abdominal segments 1-10; L, larva; P, pupa; and A, adult; max, maxillary; lab, labial; ?, not known. Homoeotic mutants are as follows. *Dachs* (*Dch*) shortens the antennae and legs, and also shows partial noncomplementation with *maxillopedia* (*mxp*). In a sample of 60 *Dch/mxp* beetles, 22 (=37%) expressed a weak *mxp* phenotype. The penetrance of homozygous *mxp* is typically 95-99%. Sterile *Dch* beetles (presumably homozygous) have normal maxillae. Mutant allele *prothoraxless* (*ptl*) is incompletely recessive and is homozygous viable and fertile in both sexes. Otherwise it conforms to the earlier description²⁴. The mutant *alate prothorax* (*apt*) is recessive, fully penetrant, and is homozygous viable and fertile. Dorsally, *apt* transforms the pronotum into a mesonotum-like structure including (in extreme cases) a recognizable mesoscutellum and median scutal suture on the middorsal posterior region of the transformed pronotum. Latera buds appear on the pronotum, but do not differentiate into elytra-like structures. Adult male *apt* homozygotes have reduced femoral sex patches, suggesting a ventral posteriorization of the prothorax as well. (Sex patches are confined to the prothoracic femurs of adult males in *T. castaneum*.) Otherwise *apt* conforms to the earlier description²⁵. Adults in *missing abdominal sternites* (*mas*) lack A1-A3 sternites²⁶. Apparently only the ventral compartments are affected, because *mas* adults have normal dorso-lateral A3 tergites. No other dorsal cuticular landmarks are known for A1-A3. The *partially pointed abdominal sternites* (*ppas*) mutant²⁷ has here been renamed *pointed abdominal sternite* (*pas*). In addition to the previously described ventral A4 → A3 (and rarely A5 → A3) transformation, the dorso-lateral A3 tergite is also reiterated on A4 in *pas* homozygotes. Also, a pair of membranous sacs develop at the posterior margin of A6 during adult eclosion. These sacs appear to be haemolymph-filled protrusions of the body wall at weak points in the intersegmental membrane in a position corresponding to that of the notches in A7. We interpret this as a recessive A6 → A7 transformation. Although the caudal segments are highly differentiated and their intersegmental boundaries are obscured, the larval and pupal appendages known as urogomphi appear to belong to A9. The *extra urogomphi* (*eu*) mutant produces an extra pair just posterior to these, and thus seems to represent an A10 → A9 transformation. In *Tetraurogomphi* (*Tu*) pupae the genital lobes (also apparently located in A9) are sometimes reiterated posteriorly²⁸. *Tu* and *labiopedia* (*lp*) are extinct.

at least one recessive, homozygous viable and fertile allele for each locus. These mutations produce segment-limited pupal or adult transformations in the antennal, maxillary or labial segments of the head, in the prothorax, or in abdominal segments 1, 2, 3, 4, 5, 6 or 10. Two of these, *maxillopedia* (*mxp*) and *missing abdominal sternites* (*mas*), are known to be closely linked on the second linkage group (LG 2)¹⁸. Two others, *antennapedia* (*ap*) and *alate prothorax* (*apt*), had been assigned to the eighth and ninth LGs, respectively^{14,15}. We showed previously that the latter assignment was incorrect, and that *apt* is actually located on LG 2, very close to *mxp* and *mas* (ref. 16 and data not shown). My preliminary linkage analysis of the remaining mutants using two- and three-point crosses indicates that six of the seven known homoeotic loci are closely linked on the second linkage group within a genetic distance of three map units and are arranged in a linear sequence identical to that of the corresponding body segments. *Antennapedia* assort independently of all other homoeotic loci. Mutant phenotypes for five of the six clustered loci are shown in Fig. 1.

To confirm this sequence, I constructed two strains homozygous for *abcd* and *cde*, respectively, where *a-e* represent recessive mutations in the five loci in anterior-to-posterior sequence, namely *mxp*, *apt*, *mas*, *pointed abdominal sternite* (*pas*) and *extra urogomphi* (*eu*). The sixth locus, *prothoraxless* (*ptl*), was not included. It is tightly linked to *apt*, but recombines with all other homoeotic loci (data not shown). I then did overlapping three- and four-factor test crosses of the type *abcd*/++++ × *abcd*/*abcd* and *cde*/+++ × *cde*/*cde* (Table 2). From this type of testcross, meiotic recombination in the heterozygous parent gives rise to a unique set of 2(*n*-1) possible crossover phenotypes (*n*-1

reciprocal pairs) for any given *n*-point sequence, neglecting multiple crossover. For testcross 1 in Table 2 (*n*=4) all six possible crossover types were recovered and the unique sequence *abcd* could thus be inferred. For testcross 2 (*n*=3) only the *eu* phenotypic class could be reliably scored, but because both the possible crossover types were recovered, the unique sequence *cde* could be inferred. By superimposing these two results the correct sequence is confirmed to be *abcde*. From these data the genetic distances between adjacent loci were calculated to be *mxp-apt*, 0.7 ± 0.1 centimorgans (cM); *apt-mas*, 1.2 ± 0.2 cM; *mas-pas*, 0.08 ± 0.03 cM; and *pas-eu*, 0.6 ± 0.2 cM.

Figure 2 summarizes the linkage relationships and segmental domains of the five loci. My working hypothesis is that this

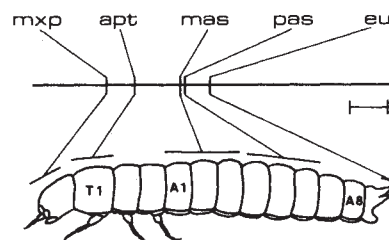


Fig. 2 Linkage relationships and segmental domains of homoeotic loci. Horizontal bar represents one centimorgan (cM) (=1% recombination) on the map of linkage group 2 (top). The first thoracic and first and eighth abdominal segments are indicated on the sketch of a last-instar larva (below).

Table 2 Determination of the linear sequence of five homoeotic loci in the homoeotic gene cluster (HOM-C) of *Tribolium castaneum*

Testcross 1		Testcross 2	
Phenotype	No. of progeny	Phenotype	No. of progeny
abcd	1,963	cde	1,303
++++	3,024	+de	2
a+++	17	++e	8
+bcd	32	c+e	0
ab++	26	Total	1,313
++cd	35		
abc+	1		
+++d	2		
all others	0		
Total	5,100		

Determination of the linear sequence of five homoeotic loci in the homoeotic gene cluster (HOM-C) of *Tribolium castaneum*. Genotypes of the form *abcde* represent the five loci in anterior-to-posterior sequence, namely *maxillopedia* (*mxp*), *alate prothorax* (*apt*), *missing abdominal sternites* (*mas*), *pointed abdominal sternite* (*pas*) and *extra urogomphi* (*eu*). Testcross 1 was *abcd/++++ × abcd/abcd*. Testcross 2 was *cde/+++ × cde/cde*. Mutation *a* (*mxp*) is only ~99% penetrant which explains the excess of +bcd phenotypes over the reciprocal class a+++, that is, half the apparent +bcd crossovers were probably nonpenetrant *abcd* noncrossovers. Mutation *e* (*eu*) is ~20–80% penetrant, depending upon temperature and genetic background. Thus, only the 'e' phenotypic class was scored for three-point linkage determination. Mutations *b*, *c* and *d* are 100% penetrant in most crosses. The unique five-point sequence predicted by the data in this table was independently confirmed by a series of two- and three-point crosses.

cluster, which I will call the homoeotic cluster or HOM-C, includes regions homologous to both the ANT-C and BX-C, but in juxtaposition rather than separated as in *Drosophila*. The HOM-C appears to span ~2.6 cM, roughly 1–2 orders of magnitude greater than the combined distances spanned by the two complexes in *Drosophila*. It may be that the HOM-C is physically much larger than the combined ANT-C and BX-C, or it may be that it is comparable in size, but that recombination is relatively more frequent in *Tribolium*. Recombination within the ANT-C is centromerically suppressed in *D. melanogaster*, but the chromosomal location of the HOM-C in *Tribolium*, as well as the overall genome size, is unknown. The HOM-C is flanked on one side (near *eu*) by *Reindeer* (*Rd*), a dominant marker located ~25–40 units away (ref. 19 and data not shown) and on the other side (near *mxp*) by *R(2)2*, a chromosome rearrangement with a breakpoint about 2–5 units away. *Pearl* (*p*) is a recessive marker located 50 units away, on the same side as *R(2)2* (ref. 16 and data not shown).

The extent of homology between the HOM-C of *Tribolium* and the ANT-C and BX-C of *Drosophila* can only be assessed through more detailed developmental and molecular studies. As expected the *T. castaneum* genome includes restriction fragments with sequence similarity to the *Drosophila* Antennapedia-type homoeobox (T. Johnson, R. Denell and R.B., unpublished data). Also, we have already isolated more than 40 new mutations in HOM-C genes and are examining their genetic relationships and developmental impact on all stages of the life cycle. These mutations define at least three lethal complementation groups, namely *maxillopedia* (*mxp*), *Cephalothorax* (*Cx*) and *Abdominal* (*Abd*). Mutations in the *maxillopedia* group cause posteriorward transformations of the antennae, or of the maxillae and labium. Mutations in *Cephalothorax* result in posteriorward transformations of the posterior head and prothorax, or errors in cephalization of the embryo. *Abdominal* mutations (including *mas* and *pas*) produce a complex variety of anteriorward and/or posteriorward transformations of abdominal segments 1–8. These new mutations will be described in detail elsewhere.

Received 13 January; accepted 13 March 1987.

- Wakimoto, B. T. & Kaufman, T. C. *Dev. Biol.* **81**, 51–64 (1981).
- Lewis, E. B. *Nature* **276**, 565–570 (1978).
- Sanchez-Herrero, E., Vernos, I., Marco, R. & Morata, G. *Nature* **313**, 108–113 (1985).
- Morata, G., Sanchez-Herrero, E. & Casanova, J. *Cell Differ.* **18**, 67–78 (1986).
- Karch, F. *et al. Cell* **43**, 81–96 (1985).
- Bender, W. *et al. Science* **221**, 23–29 (1983).
- Kaufman, T. C. in *Time, Space and Pattern in Development* (eds Jeffery, E. R. & Raff, R. A.) 365–383 (Liss, New York, 1983).
- Kaufman, T. C. & Abbot, M. K. in *Molecular Aspects of Early Development* (eds Malacinski, G. M. & Klein, W. H.) 189–218 (Plenum, New York, 1984).
- Struhl, G. *Nature* **308**, 454–457 (1984).
- Beachy, P. A., Helfand, S. L. & Hogness, D. S. *Nature* **313**, 545–551 (1985).
- Garcia-Bellido, A. *Am. Zool.* **17**, 613–629 (1977).
- Ouweneel, W. J. *Adv. Genet.* **18**, 179–248 (1976).
- Tazima, Y. *The Genetics of the Silkworm* (Logos, London, 1964).
- Sokoloff, A. *The Genetics of Tribolium and Related Species* (Academic, New York, 1966).
- Sokoloff, A. *The Biology of Tribolium* Vol. 3 (Oxford University Press, 1977).
- Beeman, R. W., Johnson, T. R. & Nanis, S. M. *J. Hered.* **77**, 451–456 (1986).
- Wool, D. & Medlinger, S. *Genetica* **44**, 496–504 (1973).
- Sokoloff, A., Ferrone, R. F., Chaney, J. D., Braden, J. & Muñoz, R. J. *Genome* **29**, 26–33 (1987).
- Dawson, P. S. & Berends, K. L. *Can. J. Genet. Cytol.* **27**, 276–278 (1985).
- Englert, D. C. & Bell, A. E. *Can. J. Genet. Cytol.* **5**, 467–471 (1963).
- Hoy, M. A. *Tribolium Inf. Bull.* **9**, 85 (1966).
- Sokoloff, A. *Tribolium Inf. Bull.* **22**, 79 (1982).
- Dawson, P. S. *J. Hered.* **59**, 188–190 (1968).
- Lasley, E. L. & Sokoloff, A. *Tribolium Inf. Bull.* **3**, 22 (1960).
- Sokoloff, A. *Tribolium Inf. Bull.* **8**, 44 (1965).
- Hoy, M. A. & Sokoloff, A. *Tribolium Inf. Bull.* **7**, 48 (1964).
- Sokoloff, A. *Tribolium Inf. Bull.* **6**, 23 (1963).
- Hoy, M. A. *Tribolium Inf. Bull.* **9**, 90 (1966).

Kinetics of smooth and skeletal muscle activation by laser pulse photolysis of caged inositol 1,4,5-trisphosphate

J. W. Walker*, A. V. Somlyo†, Y. E. Goldman†, A. P. Somlyo† & D. R. Trentham*

* National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

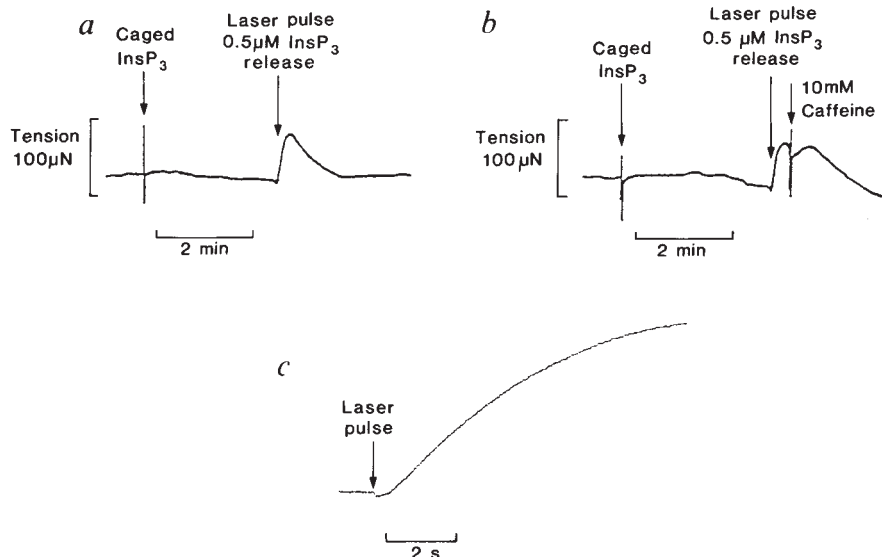
† Pennsylvania Muscle Institute and Department of Physiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6083, USA

Inositol 1,4,5-trisphosphate (InsP_3) can stimulate skinned smooth and skeletal muscle to contract by initiating Ca^{2+} release from the sarcoplasmic reticulum^{1–5,27}. Whether this process is an integral component of the *in vivo* muscle activation mechanism was tested by releasing InsP_3 rapidly within skinned muscle fibres of rabbit main pulmonary artery and frog semitendinosus. InsP_3 was liberated on laser pulse photolysis of a photolabile but biologically inactive precursor of InsP_3 termed caged InsP_3 . Caged InsP_3 is a mixture of compounds in which InsP_3 is esterified with 1(2-nitrophenyl)diazoethane (probably at the P^4 - or P^5 -position). Photochemical release of InsP_3 induced a full contraction in both muscles at physiological free Mg^{2+} concentrations, but only in the smooth muscle were the InsP_3 concentration (0.5 μM) and the activation rate compatible with the *in vivo* physiological response. Endogenous InsP_3 -specific phosphatase activity was present in smooth muscle and had about 35-fold greater activity than that in the skeletal-muscle preparation. Caged InsP_3 was not susceptible to phosphatases in either preparation.

Skeletal-muscle contraction is activated by Ca^{2+} released from the sarcoplasmic reticulum (SR) following electrical depolarization of the cell-surface membrane⁶. Smooth-muscle contraction is generally thought to be activated by myosin light-chain phosphorylation⁷ that in turn depends on the release of Ca^{2+} from the SR by electrical depolarization of the cell membrane⁸ or by pharmacomechanical coupling⁹ (following binding of neurotransmitter, independent of membrane potential). In neither case is the mechanism of signal transmission between the surface membrane and the SR well understood. Following the discovery that in non-muscle tissues inositol 1,4,5-trisphosphate (InsP_3) mediated intracellular Ca^{2+} mobilization in many

Fig. 1 Tension transients initiated by liberation of InsP_3 in smooth muscle. *a*, Chart recording of tension from a strip of rabbit main pulmonary artery (100 μm wide by 350 μm thick) permeabilized by 50 $\mu\text{g ml}^{-1}$ saponin in the presence of 0.2 mM EGTA at 20 $^{\circ}\text{C}$, pH 7.1. The SR was loaded with Ca^{2+} essentially as in ref. 1 by transferring the strip into a solution, also at 20 $^{\circ}\text{C}$, pH 7.1, of 1.5 mM free Mg^{2+} , 5.5 mM MgATP, 90 μM EGTA, 9.5 mM creatine phosphate, 50 units ml^{-1} creatine kinase, 5 μM calmodulin, 1 mM glutathione, 100 mM potassium piperazine- N,N' -bis(2-ethanesulphonate) (PIPES) and made up to an ionic strength of 0.20 M with potassium 1,6-diaminohexane- N,N,N',N' -tetraacetate (HDTA). Mitochondrial blockers, 1 mM KCN, 1 $\mu\text{g ml}^{-1}$ oligomycin and 1 μM of the protease inhibitor leupeptin were included. At the first arrow, the strip was immersed in the same solution containing 5 μM caged InsP_3 and, at the second arrow, laser pulse photolysis caused the release of 0.5 μM InsP_3 . The concentration of released InsP_3 following the laser pulse was determined by HPLC analysis of the radioactive InsP_3 and caged InsP_3 in the irradiated solution. *b*, The protocol of *a* was repeated, except that 10 mM caffeine was added. *c*, Fast recording of *a* showing response to InsP_3 release.

Methods. Caged InsP_3 (details of synthesis available on request) was prepared from the lithium salt of InsP_3 (Calbiochem) and [^3H] InsP_3 (NEN/Dupont). A mixture of singly esterified isomers (separated from the doubly and triply esterified forms) was rechromatographed by the method developed to resolve $\text{Ins}(1,4,5)\text{P}_3$ and $\text{Ins}(1,3,4)\text{P}_3$ (ref. 25). Caged InsP_3 (two of the three singly esterified isomers) was eluted ahead of the other isomer, that, like InsP_3 , induced contractions in smooth muscle and was susceptible to the endogenous smooth muscle phosphatase. We infer from these biological activities, which imply free phosphate groups at the 4- and 5-positions^{22,26}, that this other isomer is InsP_3 esterified on the P^1 -position. These results suggest that the caged InsP_3 used in the present work is a mixture of the P^4 - and P^5 -isomers. Further characterization of caged InsP_3 and related compounds was from their ion-exchange mobilities and specific activities determined from the ratio of the ^3H label to the concentration of 1(2-nitrophenyl)ethyl groups measured from $A_{265\text{nm}}$ ($\epsilon = 4,240 \text{ M}^{-1} \text{ cm}^{-1}$)¹⁴.

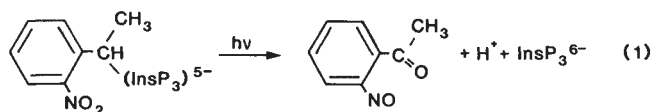


receptor-effector systems¹⁰, evidence has been obtained to support a proposed second messenger role for InsP_3 in muscle. For instance, externally applied InsP_3 has been shown to induce Ca^{2+} release from SR and to evoke contractions in smooth^{1,2}, cardiac¹¹ and skeletal³⁻⁵ muscles that have had their surface membranes permeabilized or removed (skinned). However, contractions induced by topically added InsP_3 in fast twitch skeletal muscles are slower than *in vivo* responses to physiological stimuli and, in fact, not be detected in some instances¹².

As a critical test of the second messenger hypothesis, we have made time-resolved measurements of InsP_3 -induced contractions by liberating InsP_3 rapidly in skinned fibres by laser pulse photolysis of a photolabile, but biologically inactive precursor of InsP_3 (caged InsP_3). Rate limitation owing to diffusion of InsP_3 into the fibres is thus overcome. We have also measured the endogenous phosphatase activity in the muscle preparations to determine whether caged InsP_3 is hydrolysed and to estimate how rapidly the photochemically released InsP_3 is degraded.

To investigate smooth-muscle activation, rate measurements were made in rabbit main pulmonary artery (MPA), and compared to rates of noradrenaline-stimulated contractions. For skeletal-muscle activation, frog (*Rana temporaria*) semitendinosus fibres were used and the tension responses compared to electrically stimulated contractions.

The photosensitive precursor of InsP_3 (termed caged InsP_3) is a mixture of two P-1(2-nitrophenyl)ethyl esters of InsP_3 (probably the P^4 - and P^5 -esters, see Fig. 1 legend; note the definition of caged InsP_3 differs from that used in ref. 13, where the term denoted a mixture of all three esters).



Equation (1) shows the photolysis reaction. Photolysis kinetics and the sensitivity of the muscle preparations to laser pulse

photolysis were measured as described previously^{14,15}. Following photolysis the fibres responded normally to the addition of topically added InsP_3 or caffeine, and showed no evidence of damage due to the laser pulses or the by-product, 2-nitrosoacetophenone. A single laser pulse (347 nm, 50–100 mJ) onto caged InsP_3 caused InsP_3 release with a half-time ($t_{1/2}$) of 3 ms. In the experiments that follow, laser pulse photolysis of caged InsP_3 was used to release InsP_3 uniformly and rapidly within the cytoplasmic space of the muscle fibre preparations and in the surrounding medium, and so measure the kinetics of InsP_3 -induced contractions.

First, the biological properties of caged InsP_3 were determined. Caged InsP_3 was used at a concentration below that which would induce a contraction before photolysis (Figs 1a and 2a). In the case of smooth muscle, where InsP_3 -induced contractions are graded and highly reproducible¹, caged InsP_3 at the concentrations used in photolysis experiments (5–10 μM) did not suppress or enhance contractions induced by InsP_3 . Caged InsP_3 was not susceptible to endogenous phosphatase activity (see Fig. 3 which illustrates the presence of InsP_3 phosphatases).

On photolysis of 5 μM caged InsP_3 in strips of MPA, 0.5 μM InsP_3 was generated and caused contractions to the same tension as those caused by topically added 20 μM InsP_3 or 10 mM caffeine (either during a contraction (Fig. 1b) or from the relaxed state (data not shown)). The rates of contraction induced by photo-released InsP_3 were similar to those observed in pharmacomechanical coupling: $t_{1/2}$ to peak tension was ~ 3 s and was equal to or slightly less than $t_{1/2}$ observed in smooth muscle activated by noradrenaline¹⁶. The kinetics of the main portion of the tension rise are compatible with the rate of myosin light-chain phosphorylation by ATP in the presence of myosin light-chain kinase^{17,18} (see also Fig. 1b of ref. 13). The lag phase (Fig. 1c) includes the time taken by Ca^{2+} release and Ca^{2+} -calmodulin activation of the kinase.

In skeletal muscle, the protocol (Fig. 2) was similar to that used for MPA, except that the solutions were based on those

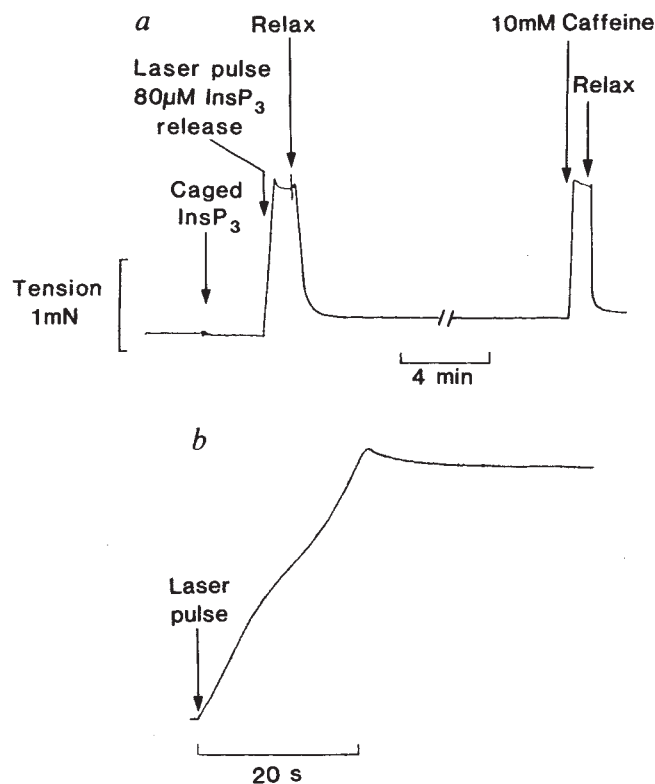


Fig. 2 Tension transients initiated by liberation of InsP_3 in skeletal muscle. *a*, Continuous chart recording of tension from a mechanically skinned single fibre from frog semitendinosus at 20°C , pH 7.1 in a solution of 0.5 mM free Mg^{2+} , 100 μM EGTA, 100 nM free Ca^{2+} , 110 mM potassium aspartate, 20 mM potassium 3(*N*-morpholino)propane sulphonate (MOPS), 5 mM creatine phosphate, 3 mM ATP, 1 mM glutathione, 1 mM KCN, 1 $\mu\text{g ml}^{-1}$ oligomycin and 1 μM leupeptin. The SR in the fibre was loaded by 5-min exposure to the above solution. At the first arrow the fibre was immersed in the same solution containing 750 μM caged InsP_3 and, at the second arrow, laser pulse photolysis released 80 μM InsP_3 . The fibre was relaxed, its SR was reloaded with Ca^{2+} as before, and the fibre was then immersed into the same solution now containing 10 mM caffeine (arrow) and finally relaxed again. *b*, Fast recording of *a* during InsP_3 release. The response of fibres to photolysis of caged InsP_3 was variable. Six fibres contracted on release of 25–80 μM InsP_3 . Three fibres were unresponsive to release of 45–50 μM InsP_3 , nor did any (of the total of 13) fibres respond to 0.5–25 μM InsP_3 . Caged InsP_3 at concentrations of 0.375–1.5 mM led to contractions with threshold varying from fibre to fibre. A contraction on photolysis of caged InsP_3 only occurred when the bathing concentration of caged InsP_3 was just below the threshold for that fibre. Prolonged incubation (20 min) in caged InsP_3 to see whether this allowed better permeation of restricted spaces did not lead to an enhanced sensitivity to photochemically liberated InsP_3 . Exposure of two fibres to saponin treatment similar to the smooth-muscle skinning procedure after mechanical skinning also did not lead to an increase in sensitivity to InsP_3 .

in ref. 3. We found that 50- to 100-fold higher concentrations of InsP_3 had to be liberated to elicit contractions, which were slow. The $t_{1/2}$ to peak tension was three orders of magnitude greater than $t_{1/2}$ associated with electrical stimulation^{19,20}. The time courses of contractions initiated by photolysis of caged InsP_3 in skeletal muscle were complex and highly variable. In some fibres tension rose immediately following the laser pulse as in Fig. 2*b*, but in other fibres a delay of up to 15 s occurred before tension developed. The $t_{1/2}$ of tension development was greater than 10 s in all cases. These results suggest that the InsP_3 -induced contractions involve a highly cooperative process

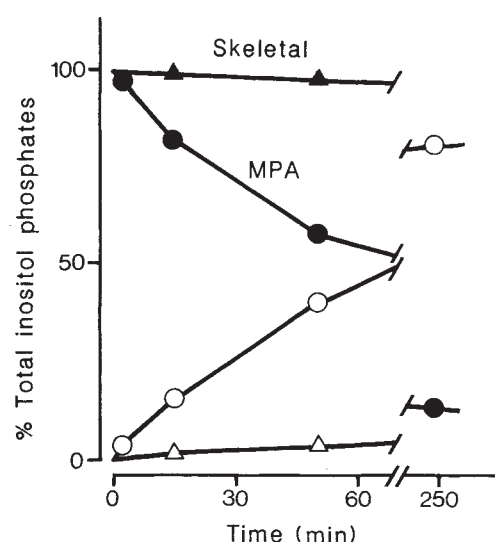


Fig. 3 InsP_3 phosphatase activity in skinned muscle fibres. MPA fibre bundles skinned with 50 $\mu\text{g ml}^{-1}$ saponin were incubated in small (45 μl) baths of buffer (see Fig. 1 legend for constituents) containing 200 μM [^3H] InsP_3 for times indicated. The InsP_3 (●) and InsP_2 (○) contents of the incubation medium were determined by HPLC and liquid scintillation counting. Data points represent mean values for three different fibre preparations of smooth muscle. The s.e.m. for each data point was within 10% of the mean value. The hydrolysis product of InsP_3 was identified as InsP_2 by its co-elution with [^3H] $\text{Ins}(1,4)\text{P}_2$ (NEN/Dupont). Further hydrolysis of InsP_2 to InsP_1 or inositol was minimal (4% of total c.p.m. over 4 h). Hydrolysis rates within the fibre were calculated over the first 15 min of the reaction, assuming that the rate was not limited by diffusion of [^3H] InsP_3 into the fibre. Fibre volumes were $\sim 1/500$ of the trough volume. The experimental protocol was repeated on mechanically and saponin-skinned skeletal fibres (▲, InsP_3 ; △, InsP_2). The 1.5% hydrolysis of InsP_3 at 50 min had an s.e.m. of 0.40%. In control experiments, no additional phosphatase activity was detected when incubations were performed in the saponin-skipping trough with the fibre present, and no phosphatase activity was detected in the skinning trough after the fibre was removed. Thus, no evidence was obtained in either muscle preparation for soluble phosphatase that diffused out of fibres during skinning. Also saponin did not inhibit the measured InsP_3 phosphatase activity. When free Mg^{2+} concentration was reduced from 1.5 to 0.04 mM, the smooth muscle InsP_3 phosphatase activity decreased to $\sim 40\%$ of its previous value. The human erythrocyte enzyme displays a similar Mg^{2+} dependence²².

such as the well-known phenomenon of Ca^{2+} -induced release of Ca^{2+} (ref. 21).

In earlier work^{3,4} it was often necessary to lower the free Mg^{2+} concentration below the physiological level to 40 μM to obtain reproducible contractions of skeletal muscle fibres at low InsP_3 concentrations (30 μM). This manipulation was intended to inhibit InsP_3 hydrolysis²², but those results should be interpreted with caution because low Mg^{2+} concentration also enhances Ca^{2+} -induced release of Ca^{2+} (ref. 21). The contractions of Fig. 2 were initiated by photolysis of caged InsP_3 at 0.5 mM free Mg^{2+} which is in the physiological range. The extent and rate of tension development were similar at 40 μM free Mg^{2+} .

Cells that use InsP_3 as an intracellular messenger possess an enzyme, InsP_3 5-phosphatase, that inactivates InsP_3 by conversion to inositol 1,4-bisphosphate (InsP_2)¹⁰. We have measured InsP_3 phosphatase activity in our skinned muscle preparations (Fig. 3). InsP_3 phosphatase activity in MPA corresponded to hydrolysis of 200 μM InsP_3 at 15–20 $\mu\text{M s}^{-1}$ within the fibres, when allowance was made for the fibre volume relative to that

of the bathing solution. This high level of activity in MPA is consistent with the InsP_3 messenger hypothesis for pharmacomechanical coupling. Similar results (not shown) were obtained with rabbit portal mesenteric vein. InsP_3 formed during the photolysis experiment may be degraded at less than $15\text{--}20\ \mu\text{M s}^{-1}$ due to a relatively high K_m of the enzyme for InsP_3 . For instance a K_m of $25\ \mu\text{M}$, which is the measured value for the human erythrocyte enzyme²², would lower the initial rate of hydrolysis of $0.5\ \mu\text{M}$ InsP_3 to $0.3\ \mu\text{M s}^{-1}$.

In frog skeletal fibres (mechanically or saponin skinned) the InsP_3 hydrolysis rate indicated that the phosphatase activity was $0.5\ \mu\text{M s}^{-1}$ within the fibres which is 35-fold less than that in MPA (Fig. 3). The InsP_3 phosphatase activity of mechanically isolated (everted) sarcolemma was also low (that is, the measured activity would have caused hydrolysis of InsP_3 at about $0.5\ \mu\text{M s}^{-1}$ within the fibre volume). This experiment and the results from saponin-skinned skeletal muscle fibres excluded the possibility of high phosphatase activity in the surface membrane. The slow InsP_3 hydrolysis rate in skinned skeletal muscle does not support the hypothesis^{3,4} that slow responses and apparent insensitivity of fibres to topically added InsP_3 are the result of a high InsP_3 -phosphatase activity.

In conclusion, our results provide strong support for the hypothesis that InsP_3 is a second messenger in the regulation of smooth-muscle contraction. Besides establishing the kinetics of InsP_3 -induced contraction in smooth muscle, the method of laser pulse photolysis of caged InsP_3 reveals that Ca^{2+} release is an order of magnitude more sensitive to InsP_3 than previously measured and circumvents the problem of rapid breakdown of InsP_3 by endogenous phosphatase activity. However, our results and those of others¹² do not favour the hypothesis for fast-twitch skeletal muscle. We cannot exclude the possibility that an essential cofactor required for rapid InsP_3 action is lost in skinned skeletal muscle fibres. Alternatively, InsP_3 may modulate excitation-contraction coupling in skeletal muscle²³ without being the primary messenger that delivers the message across the gap between the T-tubule and the sarcoplasmic reticulum²⁴.

This work was supported by an NIH grant to the Pennsylvania Muscle Institute and by the MRC of the UK.

Received 22 December 1986; accepted 2 April 1987.

- Somlyo, A. V., Bond, M., Somlyo, A. P. & Scarpa, A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5231–5235 (1985).
- Suematsu, E., Hirata, M., Hashimoto, T. & Kuriyama, H. *Biochem. biophys. Res. Commun.* **120**, 481–485 (1984).
- Vergara, J., Tsien, R. Y. & Delay, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6352–6356 (1985).
- Volpe, P., Salvati, G., Di Virgilio, F. & Pozzan, T. *Nature* **316**, 347–349 (1985).
- Donaldson, S. K., Goldberg, N. D., Walseth, T. F. & Huetteman, D. A. *Biophys. J.* **49**, 191a (1986).
- Baylor, S. M. in *The Handbook of Physiology: Skeletal Muscle* (eds Peachey, L. D., Adrian, R. H. & Geiger, S. R.) 355–379 (Am. Physiol. Soc., Bethesda, Maryland, 1983).
- Hartshorne, D. J. & Gorecka, A. in *The Handbook of Physiology: The Cardiovascular System, Vascular Smooth Muscle Vol. II* (eds Bohr, D. F., Somlyo, A. P. & Sparks, H. V.) 93–120 (Am. Physiol. Soc., Bethesda, Maryland, 1980).
- Johansson, B. & Somlyo, A. P. in *The Handbook of Physiology: The Cardiovascular System, Vascular Smooth Muscle Vol. II* (eds Bohr, D. F., Somlyo, A. P. & Sparks, H. V.) 301–324 (Am. Physiol. Soc., Bethesda, Maryland, 1980).
- Somlyo, A. P. & Somlyo, A. V. *J. Pharmac. exp. Ther.* **159**, 129–145 (1968).
- Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315–321 (1984).
- Fabiato, A. in *The Regulation of Excitation and Contraction* (ed. Nathan, R. D.) 283–293 (Academic, New York, 1986).
- Lea, T. J., Griffiths, P. J., Tregear, R. T. & Ashley, C. C. *FEBS Lett.* **207**, 153–161 (1986).
- Goldman, Y. E., Reid, G. P., Somlyo, A. P., Somlyo, A. V., Trentham, D. R. & Walker, J. W. *J. Physiol., Lond.* **377**, 100P (1986).
- McCray, J. A., Herbet, L., Kihara, T. & Trentham, D. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7237–7241 (1980).
- Goldman, Y. E., Hibberd, M. G. & Trentham, D. R. *J. Physiol., Lond.* **354**, 577–604 (1984).
- Devine, C. E., Somlyo, A. V. & Somlyo, A. P. *J. cell Biol.* **52**, 690–718 (1972).
- Kamm, K. E. & Stull, J. T. *Science* **232**, 80–82 (1986).
- Somlyo, A. P. *et al.* *J. Mus. Res. Cell Motil.* **7**, 38 (1986).
- Kress, M., Huxley, H. E., Faruqi, A. R. & Hendrix, J. J. *molec. Biol.* **188**, 325–342 (1986).
- Stephenson, E. W. *J. gen. Physiol.* **77**, 419–443 (1981).
- Endo, M. *Physiol. Rev.* **77**, 171–188 (1971).
- Downes, C. P., Mussat, M. C. & Michell, R. H. *Biochem. J.* **203**, 169–177 (1983).
- Volpe, P., Di Virgilio, F., Pozzan, T. & Salvati, G. *FEBS Lett.* **97**, 1–4 (1986).
- Somlyo, A. P. *Nature* **316**, 298–299 (1985).
- Irvine, R. F., Anggard, E. E., Letcher, A. J. & Downes, C. P. *Biochem. J.* **229**, 505–511 (1985).
- Irvine, R. F., Brown, K. D. & Berridge, M. J. *Biochem. J.* **222**, 269–272 (1984).
- Iino, M. *Biochem. biophys. Res. Commun.* **142**, 47–52 (1987).

Interactions between DNA and coat protein in the structure and assembly of filamentous bacteriophage fd

Gary J. Hunter, David H. Rowitch & Richard N. Perham

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Bacteriophage fd is a class I filamentous virus (others are M13 and f1) that comprises a circular, single-stranded DNA molecule enclosed in a cylindrical protein sheath to form a flexible particle ~890 nm long and 7 nm in diameter (for reviews, see refs 1 and 2). The viral DNA contains 6,408 nucleotides^{3–5} incorporating 10 genes, and the protein sheath is composed of about 2,700 major coat protein subunits⁶ in a shingled helical array, the symmetry of which is defined by a fivefold rotational axis combined with a twofold screw axis of pitch 3.2 nm (refs 7–9). The DNA extends throughout the length of the particle but is not base-paired and has a symmetry different from that of the protein helix. How the DNA is packed remains unclear but the number (2.4) of nucleotides packaged per major coat protein subunit is certainly not integral^{6,7}, in contrast with, say, the packaging of RNA in tobacco mosaic virus¹⁰. The coat protein subunit is 50 amino-acid residues in length and, in the virus particle, adopts a largely α -helical conformation, with the long axis of the helix aligned close to the long axis of the filament^{7–9,11}. This protein is arranged with its negatively charged N-terminal region on the outside of the filament and its positively charged C-terminal region on the inside abutting the DNA^{7,12}. We report here that positive charge on one of the four lysine side chains in the latter region has a direct effect on DNA packaging, because when this charge is absent, elongated particles are produced with lengths that can be correlated with the residual positive charge in the C-terminal region of the coat protein subunit.

The viral DNA is specifically oriented within the bacteriophage particle to accommodate a 78-nucleotide hairpin loop at one end of the virion¹³. This end of the particle also carries approximately five copies of each of two minor coat proteins, gVIIp and gIXp, encoded by bacteriophage genes VII and IX, respectively^{14–16}. The major coat protein gVIIIp, is encoded by bacteriophage gene VIII and is synthesized as a procoat molecule containing an N-terminal signal peptide of 23 amino acids (Fig. 1). The procoat molecule is processed by signal peptidase and the mature coat protein is left spanning the bacterial inner membrane with its N-terminus exposed on the outer face of the membrane and its C-terminus in the cytoplasm^{17,18}. During assembly of the virion at the bacterial inner membrane, the major coat protein is drawn from the membrane and added to the elongating particle which is being extruded. Assembly begins at the gVIIp-gIXp end of the filament^{12,19} and is probably initiated by interaction of these and other proteins in the membrane with the hairpin loop of DNA found at that end of the finished particle¹. Assembly is completed by the addition of approximately five copies of each of two other minor coat proteins, gIIp and gVp, encoded by bacteriophage genes III and VI, respectively, to the opposite end of the filament^{14,15} as it leaves the membrane.

There are four lysine residues in the C-terminal region of the major coat protein (Fig. 1a) and it is reasonable to suppose that some or all of these have an important function in neutralizing the negatively charged phosphodiester links in the encapsidated DNA⁷. The single-stranded DNA bacteriophages are favoured vehicles for site-directed mutagenesis of cloned genes²⁰. It was thought likely, therefore, that the same technique could be used to produce selected mutations directly in bacteriophage fd gene VIII, although it was recognized that only those mutations that generated a functional gVIIIp, compatible with virus assembly, would be picked up.

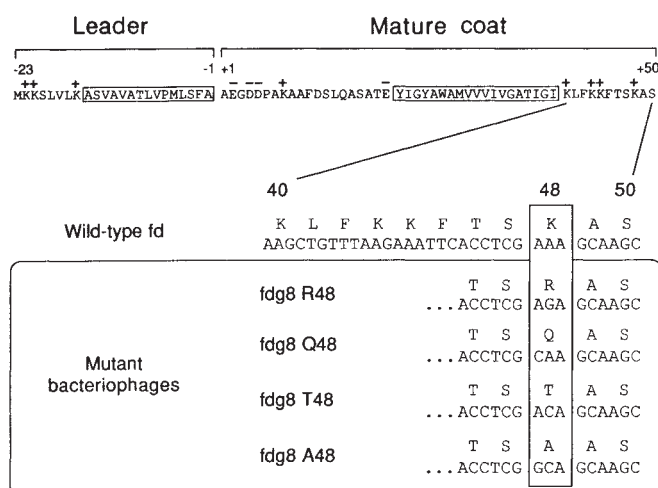


Fig. 1 *a*, Amino-acid sequence of the procoat of bacteriophage fd. Putative membrane-spanning regions are indicated by a box. Positively and negatively charged side chains are indicated by + and -, respectively. *b*, Amino-acid sequences of the C-terminal regions of the major coat proteins of bacteriophage fd and selected mutants, with partial nucleotide sequences of the corresponding genes. Shaded box, altered codons.

Methods. Mutant bacteriophages were generated by oligonucleotide-directed mutagenesis²⁶. The oligonucleotides d(AGCTTGCTCTCGAGGTGAA), d(AGCTTGCTTGCGAGGTGAA) and d(AGCTTGCTGTCGAGGTGAA) were synthesized²⁷ and hybridized to wild-type viral ssDNA to generate mutants fdg8R48, fdg8Q48 and fdg8T48, respectively. Mutant fdg8A48 was isolated as an apparent pseudorevertant from a preparation of mutant fdg8T48. The complete nucleotide sequence of gene VIII in each mutant was checked by dideoxy sequencing^{28,29}, using the synthetic oligonucleotide d(CACGTTGAAAATCTCCA) as a primer.

Mutations were designed to change the codon at position 48 of the fd coat protein gene VIII from AAA (lysine, wild-type) to AGA (arginine, fdg8R48), CAA (glutamine, fdg8Q48) and ACA (threonine, fdg8T48), respectively (Fig. 1*b*). A mutant bacteriophage carrying the desired modification to gene VIII was isolated in each instance and, from preparations of the mutant bacteriophage fdg8T48, an additional unexpected mutant (fdg8A48) was isolated that had incorporated alanine instead of threonine at position 48 (Fig. 1*b*). Because the major coat protein forms 88% by weight of the bacteriophage particle⁶, it was also possible to obtain direct confirmation of the expected change in amino-acid sequence by analysis of acid hydrolysates of purified bacteriophage samples (data not shown).

The generation of viable bacteriophage particles was in itself evidence that the modifications of the major coat protein had not totally impaired its function. To test whether the modifications had led to any significant changes in the structure of the virion, however, mutant bacteriophages were submitted to electrophoresis in agarose gels (Fig. 2*a*). This revealed that the mutant bacteriophages all had mobilities different from that of the parent fd. The changes in electrophoretic mobility could not be directly correlated with the loss (in mutants fdg8Q48, fdg8T48, fdg8A48) or retention (in fdg8R48) of the positive charge on the side chain of the amino acid occupying position 48 in the major coat protein. Electrophoretic analysis of the single-stranded DNAs (ssDNAs) extracted from the various bacteriophages (Fig. 2*b*) confirmed that none of these changes could be attributed to unsuspected changes in the length of the DNA molecules.

Charge changes in the N-terminal region of the major coat protein induce corresponding changes in electrophoretic mobility of the intact virion²¹. The C-terminal region of the major coat protein, however, is tucked away inside the bacteriophage particle abutting the DNA^{7,12}. The most plausible explanation

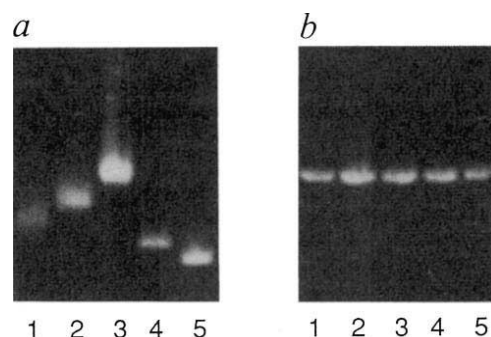


Fig. 2 Agarose gel electrophoresis of bacteriophage particles and DNA. *a*, Electrophoresis of intact virions: lanes 1-5, bacteriophages fd, fdg8R48, fdg8Q48, fdg8T48 and fdg8A48, respectively. *b*, Electrophoresis of viral ssDNA: lanes 1-5, DNA from bacteriophages fd, fdg8R48, fdg8Q48, fdg8T48 and fdg8A48, respectively.

Methods. Electrophoresis of intact virions was carried out in 2% agarose gels in 0.37 M Tris-glycine buffer, pH 9.5, and the positions of the virions were revealed by staining the viral DNA with ethidium bromide³⁰. Viral ssDNA, obtained by phenol extraction of the virions, was submitted to electrophoresis in 0.7% agarose gels in 0.04 M Tris-acetate buffer containing 2 mM EDTA, pH 7.8 and stained with ethidium bromide, using standard techniques³¹.

of the observed changes in electrophoretic mobility is that the mutations effected in this region of the polypeptide chain have caused changes in DNA-protein or protein-protein interactions in the virus particle. These might in turn have generated different surface charges on the virion or, more likely, have affected the flexibility of the filament and thus its freedom of passage during electrophoresis in agarose gels.

Dramatic evidence of structural change in the virion ensuing from mutation in the major coat protein was obtained from electron microscopy (Fig. 3). The contour lengths of the particles were measured and the results are shown in Table 1. Each virus sample examined was, within the limits of experimental error, uniform in length, and the lengths of bacteriophages fd and fdg8R48 were essentially identical. In contrast, the mutants fdg8A48, fdg8Q48 and fdg8T48 were all markedly longer, by about 35%.

Previous studies have shown that bacteriophage fd can be shortened by deleting segments of the viral DNA²², or lengthened by cloning extra DNA into the bacteriophage genome²³. The number of nucleotides packed per major coat protein subunit remains the same and the length of the particle is proportional to the length of the DNA encapsidated. This is of course one of the major attractions of the virus as a cloning vehicle.

In our experiments the length of the DNA was unchanged. The conversion of lysine 48 to arginine 48 in the major coat protein (Fig. 1) was designed to ensure that four positively charged side chains were retained in the C-terminal region of the protein close to the DNA in bacteriophage fdg8R48. No significant change in particle length was observed. The conversion of lysine 48 to alanine 48, glutamine 48 or threonine 48, on the other hand, meant that the number of such positive charges fell to three. This region of the polypeptide chain is not intimately involved in the protein-protein contacts that generate the protein capsid. Therefore, assuming that the protein helix did not change substantially, a simple explanation of the increase in length of the resulting particles (Fig. 3 and Table 1) would be that more protein subunits were required to pack the same amount of DNA. This would be independent of the nucleotide sequence of the DNA. The agreement may be fortuitous, but if each of these four lysine residues in the wild-type coat protein were contributing equally to the neutralization of the negative charges of the packaged DNA, the required increase in length of the mutated bacteriophages would be 33%.

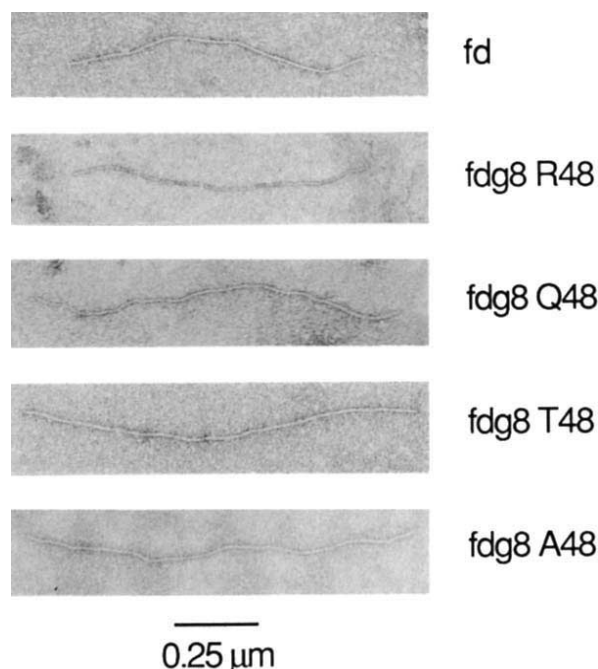


Fig. 3 Electron microscopy of bacteriophage particles. The virions were spread on copper grids in 50 mM ammonium acetate buffer, pH 5.5, and negatively stained with 2% uranyl acetate. Electron micrographs were taken on an AE1801S electron microscope, at a magnification of $\times 40,000$.

From the electron micrographs we cannot be certain that the symmetry of the coat protein helix has remained undisturbed in any or all of the mutants that we have constructed. It should be possible to tell this from X-ray diffraction analysis⁷⁻⁹. Such analysis should enable us to be more precise about the arrangement of the DNA in the mutated bacteriophages and may help us to interpret the suspected changes in flexibility of the virus particles.

Genetic reconstruction experiments have suggested that a cluster of positive charges in the C-terminal region of the major coat protein is of major importance for correct membrane insertion and processing²⁴. Our experiments demonstrate that lysine 48 is not essential for this, focusing attention on the lysine residues at positions 40, 43 and 44. Thus it may be significant that we have so far been unable to isolate a viable bacteriophage fd in which any of these three lysine residues has been mutated to an amino acid lacking a positively charged side chain. Further experiments of the kind described here and elsewhere²⁵ should throw light on this intriguing problem, as well as on the mechanism of DNA packaging in these important viruses.

We are grateful to the MRC and the Commission of the European Communities for financial support. D.H.R. thanks the Oliver Gatty Fund for the award of a Studentship. We thank

Table 1 Lengths of bacteriophage particles (nm)

Bacteriophage	Contour length $\pm$ s.d. (N)	Length relative to bacteriophage fd
fd	866 $\pm$ 37 (59)	1.00
fdg8R48	891 $\pm$ 32 (34)	1.03
fdg8Q48	1153 $\pm$ 55 (34)	1.33
fdg8T48	1196 $\pm$ 72 (36)	1.38
fdg8A48	1178 $\pm$ 57 (34)	1.36

Contour lengths of separate bacteriophage particles were measured directly from electron micrographs using a Graf/Bar sonic digitizer. Standard deviation (s.d.) and number of particles measured (N) are shown for each bacteriophage sample.

Dr D. A. Marvin for valuable discussion and Mr J. G. Lay for skilled technical assistance with the electron microscopy.

Received 26 February; accepted 31 March 1987.

- Webster, R. E. & Lopez, J. in *Virus Structure and Assembly* (ed. Casjens, S.) 235-267 (Jones and Bartlett, Boston, 1985).
- Makowski, L. in *Biological Macromolecules and Assemblies* Vol. 1 *Virus Structures* (eds Jurnak, F. A. & McPherson, A.) 203-253 (Wiley, New York, 1985).
- Beck, E. *et al. Nucleic Acids Res.* **5**, 4495-4503 (1978).
- Van Wezenbeck, P. M. G. F., Hulsebos, T. J. M. & Schoenmakers, J. G. G. *Gene* **11**, 129-148 (1980).
- Hill, D. F. & Petersen, G. B. *J. Virol.* **44**, 32-46 (1982).
- Day, L. A. & Wiseman, R. L. in *The Single-Stranded DNA Phages* (eds Denhardt, D. T., Dressler, D. & Ray, D. S.) 605-625 (Cold Spring Harbor Laboratory, New York, 1978).
- Marvin, D. A. in *The Single-Stranded DNA Phages* (eds Denhardt, D. T., Dressler, D. & Ray, D. S.) 583-603 (Cold Spring Harbor Laboratory, New York, 1978).
- Banner, D. W., Nave, C. & Marvin, D. A. *Nature* **289**, 814-816 (1981).
- Makowski, L. & Caspar, D. L. D. *J. molec. Biol.* **145**, 611-617 (1981).
- Holmes, K. C. in *Structural Molecular Biology* (eds Davies, D. B., Saenger, W. & Danyluk, S. S.) 475-505 (Plenum, New York, 1982).
- Cross, T. A., Tsang, P. & Opella, S. J. *Biochemistry* **22**, 721-726 (1983).
- Armstrong, J., Hewitt, J. A. & Perham, R. N. *EMBO J.* **2**, 1641-1646 (1983).
- Ikoku, A. S. & Hearst, J. E. *J. molec. Biol.* **151**, 245-259 (1981).
- Grant, R. A., Lin, T.-C., Konigsberg, W. & Webster, R. E. *J. biol. Chem.* **256**, 539-546 (1981).
- Simons, G. F. M., Konings, R. N. H. & Schoenmakers, J. G. G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4194-4198 (1981).
- Webster, R. E., Grant, R. A. & Hamilton, L. A. W. *J. molec. Biol.* **152**, 357-374 (1981).
- Wickner, W. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4749-4753 (1975).
- Ohkawa, I. & Webster, R. E. *J. biol. Chem.* **256**, 9951-9958 (1981).
- Lopez, J. & Webster, R. E. *Virology* **127**, 177-193 (1983).
- Carter, P. *Biochem. J.* **237**, 1-7 (1986).
- Boeke, J. D., Russel, M. & Model, P. *J. molec. Biol.* **144**, 103-116 (1980).
- Griffith, J. & Kornberg, A. *Virology* **59**, 139-152 (1974).
- Herrmann, R., Neugebauer, K., Zentgraf, H. & Schaller, H. *Molec. gen. Genet.* **159**, 171-178 (1978).
- Kuhn, A., Wickner, W. & Kreil, G. *Nature* **322**, 335-339 (1986).
- Rowitch, D. H. & Perham, R. N. *J. molec. Biol.* (in the press).
- Zoller, M. J. & Smith, M. *Meth. Enzym.* **100**, 468-500 (1983).
- Mathes, H. W. D. *et al. EMBO J.* **3**, 801-805 (1984).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. J. *J. molec. Biol.* **143**, 161-178 (1980).
- Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).
- Moses, P. B. & Horiuchi, K. *Virology* **119**, 231-234 (1982).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnType ☐ OCLC ILL Subsystem

☐ Other (please specify) _____

☐ I am interested in sending my order by mail.

☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Cryofixation taking on a new look

S.A. Livesey and J.G. Linner

Probing for electron microscopy and microanalysis requires samples with undamaged ultrastructure. A new technique for nondestructive freezing has been developed for better sample preparation.

COMBINING the structural resolution of the electron microscope with direct chemical analysis has long been the goal of cell biologists. The development of probing techniques such as colloidal gold immunolabelling, autoradiography, microanalysis and hybridization histochemistry for use in electron microscopy has provided the tools to make this integration possible. This expanding list of techniques now allows the sensitive detection of proteins, elements, nucleotides and drugs, thereby greatly expanding the uses of the electron microscope for investigations in cell physiology, biochemistry, pathology and pharmacology.

However, the conventional methods of specimen preparation, employing chemical fixation, solvent dehydration and resin embedding, have provided sub-optimal samples for the application of these probes. The limitations imposed by these standard sample preparation techniques result from the combined effects of chemical crosslinking, aqueous extraction following the destruction of permeability barriers by the fixative, and solvent extraction during dehydration¹. The resulting structural alteration, redistribution and loss of soluble cellular components often leads to non-detection, or worse, artefactual cellular localization of the probe target².

For a new, more rigorous method of specimen preparation to be useful, it must minimize the effects of chemical fixation and produce a sample which is stable during the manipulations required for the application of probes. It must also be able to withstand the effects of the vacuum and electron beam in the microscope. The process described here combines cryofixation, molecular distillation drying and direct resin infiltration under vacuum³. It does not involve cryoprotection, chemical prefixation, aqueous or solvent extraction, and therefore avoids the pitfalls associated with these techniques. The end product is a stable, resin-embedded sample which is amenable to conventional sectioning and analysis.

The most common problem to be overcome in using this process is the formation of ice crystals, with concomitant destruction of the sample. Ice crystal formation must be prevented both during cryofixation of the sample and during the subsequent drying process. Cultured cells provide an ideal model system for the application of this method because they pro-

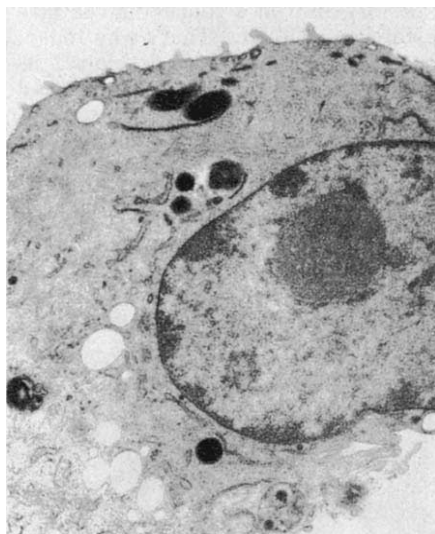


Fig. 1 Human breast cancer cell (MCF 7) grown on Thermanox, cryofixed, molecular distillation dried, vapour osmicated and embedded in Spurr's resin.

vide uniform sample parameters of size, hydration state and orientation, and are relatively easy to prepare and manipulate.

The big chill

Ice crystal formation during cryofixation can be minimized by two approaches: cryoprotection and controlled rate freezing, and ultrarapid cooling of the native sample. In view of the potential artefacts induced by cryoprotectants⁴, we have employed a method of ultrarapid cooling using the cold metal-mirror technique.

While cold metal-mirror freezing is generally accepted to be the optimum method for achieving the highest cooling rates⁵, its application to cultured cells has been limited by the practical problems of sample delivery, compression, transfer and manipulation under liquid nitrogen. In this process, cells are plated on 3 × 4 mm squares of Thermanox and maintained in culture until immediately prior to cryofixation. The design of the sample holder to carry the cells to the precooled copper block is critical, if compression damage is to be avoided. The sample holder used consists of a multilayered system composed of a metal disk attached to a magnetic plunger, followed sequentially by 1 cm soft foam, 2 mm dense foam, 2 mm 2.5 per cent agar and finally a thin brass plate onto which the Thermanox with the plated cells is placed and held in position by capillary action. Excess medium is removed by wicking with filter

paper just before cryofixation. The mirror-finished copper bar is cooled by flow-through of liquid helium and is maintained free of contaminants during the cooling process by a surrounding ultra-high hydrocarbon-free vacuum. At the time of sample delivery, the vacuum system is isolated and the vacuum around the bar reversed by the introduction of dry helium gas. The sample is applied directly to the surface of the copper bar by an electromagnetic or pneumatic shutter-plunger system.

Once cryofixed, the sample is transferred directly to liquid nitrogen and maintained at -196°C until the molecular distillation drying process is commenced. The structural integrity of the resulting sample is shown in Fig. 1.

Preliminary data using differential scanning calorimetry has confirmed a detectable phase change from amorphous to cubic ice in samples prepared this way (J. Baust, State University of New York-Binghamton, personal communication). This indicates that at least a portion of the cell water is obtained in the vitreous state when utilizing this ultrarapid cooling method on non-cryoprotected biological samples.

Distillation drying

The amorphous or microcrystalline state of cellular water is unstable, in that temperature elevation will result in water molecule migration with subsequent ice crystal formation and growth. The temperature at which this occurs in cells remains uncertain. With glycerol-protected muscle samples, the phase transition assessed by X-ray diffractograms indicates that devitrification occurs at -120°C⁶. Preliminary results with differential scanning calorimetry indicate that, rather than being a discrete event, recrystallization may occur over a relatively broad temperature range, and may be a multi-component process, beginning at -150°C and ending at -80°C (J. Baust, State University of New York-Binghamton, personal communication).

To avoid recrystallization during the removal of water from the cryofixed sample, a molecular distillation dryer is used. The samples are loaded under liquid nitrogen into a precooled copper sample holder equipped with a thermocouple and a cartridge-type heater. The loaded sample holder is then transferred to a liquid nitrogen-filled stainless steel

sample chamber which is immersed in a Dewar containing liquid nitrogen. Following electrical coupling of the sample holder, the chamber is bolted to the Dewar, the liquid nitrogen is removed by a roughing pump, and the samples are exposed to an ultrahigh vacuum generated by a helium cryopump. The equilibrium temperature of this apparatus is -177°C , well below the recrystallization temperature⁷, and the vacuum is 10^{-7} to 10^{-8} mbar.

Heating of the sample during the drying cycle is controlled by a programmable microprocessor to allow continuous incremental temperature elevation of 1°C per hour during the most critical drying phase. The controlled continuous heating, constant ultrahigh vacuum and efficient condenser surfaces (liquid nitrogen-cooled chamber wall and helium cryopump) are features essential to the successful operation of the dryer.

Once the sample is dry, the chamber wall temperature is elevated by slow evaporation of the surrounding liquid nitrogen. The previously condensed water is trapped by the helium cryopump, which is isolated subsequently by a gate valve. Depending upon the application, samples can be unloaded without further treatment, infiltrated directly with resin at low temperature, or vapour-fixed in the dry state followed by direct resin infiltration without breaking the vacuum.

Resin-embedded samples, maintained on Thermanox throughout cryofixation, drying, resin embedding and sectioning steps, can then be probed, using either a post-embedding technique⁸ or with pre-embedding fusion of the probe as in autoradiography. Elemental analysis can be performed on dried cells grown and processed on Formvar-coated grids or on resin-embedded sections (preferably non-ossicated and embedded in Lowicryl resin at low temperatures)⁹. Although at an early stage of development, cryofixation of untreated cells followed by molecular distillation drying may have importance to cultured cell preservation and storage, as well as ultrastructural and biochemical probing. □

Stephen A. Livesey and John G. Linner are at the Cryobiology Research Center, 3606A Suite 1, Research Forest Drive, The Woodlands, Texas 77380, USA. For more information, fill in reader service number 100.

1. Bullock, G.R. *J. Microsc.* **133**, 1–15 (1984).
2. Hall, T.A. *J. Microsc.* **117**, 145–163 (1984).
3. Skaer, H. *J. Microsc.* **125**, 137–147 (1981).
4. Linner, J.G., Bennett, S.C., Harrison, D.S. & Steiner, A.L. in *Science of Biological Specimen Preparation* (eds Revel, J.P., Barnard, T. & Harris, G.H.) 165–174 (Scanning Electron Microscopy, Inc., AMF O'Hare, Ill., 1986).
5. Jones, G.J. *J. Microsc.* **136**, 349–353 (1984).
6. Dubochet, J. & McDowell, A.W. in *Science of Biological Specimen Preparation* (eds Revel, J.P., Barnard, T. & Harris, G.H.) 147–152 (Scanning Electron Microscopy, Inc., AMF O'Hare Ill., 1984).
7. Rasmussen, D.H. & Mackenzie, A.P. *J. phys. Chem.* **75**, 967–971 (1971).
8. Linner, J.G., Livesey, S.A., Harrison, D.S. & Steiner, A.L. *J. Histochem. Cytochem.* **34**, 1123–1135 (1986).
9. Wroblewski, J. & Wroblewski, R. *J. Microsc.* **142**, 351–362 (1985).

Tissue culture tools

Tissue culturists of every type and description will be in Washington, D.C. for next week's US Tissue Culture Association meeting on 27–30 May. Below are some of the products they will see on a stroll through the exhibit hall.

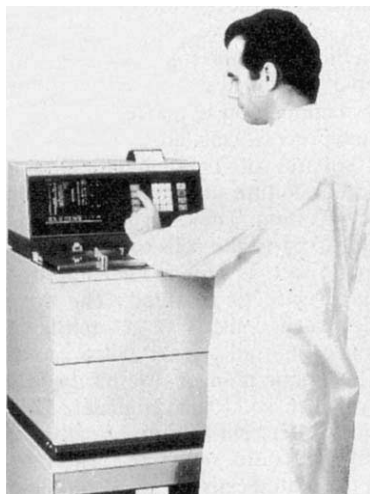
CENTRIFUGING in a cold room can give researchers cold feet. That's why Jouan, Inc., in booth 304, has developed its model MR1411 **refrigerated high-speed microcentrifuge** (Reader Service No. 101). The Jouan microcentrifuge accommodates tubes over a wide range of sizes — as small as 0.5 ml or as large as 10 ml — and spins at a maximum of 12,000 r.p.m., for a force of 13,690g. Standard features include closed-loop speed control for



Jouan's MR1411 is made for spinning down small volumes at low temperatures.

stable runs, digital speed display, and automatic dynamic electronic breaking for smooth stops without sample resuspension. The \$2,400 (US) MR1411 has a full range of optional rotors, including a drum rotor and an angle rotor. The drum rotor holds racks of different-sized tubes in a horizontal position for more even pellets, while the angle rotor is meant for large volume tubes.

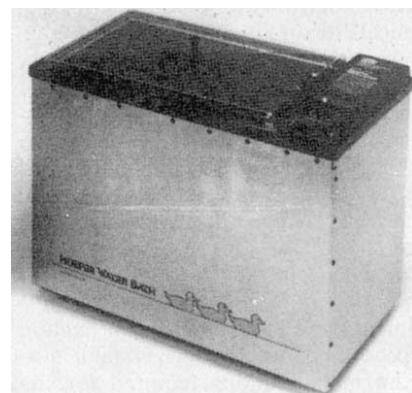
The new **tissue processors** from Fisher Scientific shield operators from exposure to the noxious fumes of reagents and paraffins fused for fixing, dehydrating and infiltrating tissue samples for microscopy



Fisher's tissue processor can be stacked, as shown, or used on the benchtop.

(Reader Service No. 102). The \$17,000 (US) Histomatic Modular Vacuum Processor (MVP) processes up to 150 tissue cassettes at a time, and the \$23,500 (US) MVP II can handle up to 300. Both units have a 64K computer memory, and an RS232 interface to allow them to be hooked up to a personal computer. Solvents for fixing, dehydrating and clearing stages are in ten plug-in containers, and media for infiltration is kept in four independently programmable paraffin compartments. Quality of reagents is analysed continuously and displayed on the CRT screen. All parameters are programmed from the keyboard, including stirrer speed, vacuum, pressure, time and temperature. Fisher will be showing off its wares in booths 205–207, 305 and 306.

In booth 502, Hoefer Scientific will be talking about its new **extra-deep water-bath**, the PR 840 (Reader Service No. 103). The \$1,600 (US) Hoefer bath



Hoefer's water bath is made to accommodate large vessels for incubation.

measures 32 cm long × 22 cm wide × 27 cm deep, permitting incubation of laboratory vessels in water 24 cm deep. The water temperature is regulated to within $\pm 0.25^{\circ}\text{C}$, and is displayed on an LCD readout. The bath's recirculation pump turns over the entire water content of the bath every three minutes, according to Hoefer, and the insulated sides and bottom of the bath minimize heat loss.

Thermolyne's new **Locator cryobiological storage systems** will be on display in booths 406 and 503 (Reader Service No. 104). Designed to help researchers put their hands on the sample they want without having to search through an entire inventory of samples, the Locator systems use numbered racks and Nalge boxes specially designed to withstand the

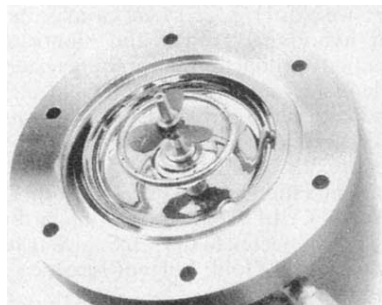


The Locator cryobiological storage systems come in two configurations to address different cold storage needs.

intense cold of liquid nitrogen. Each Nalge box has a fully numbered grid system and one-way lid to ensure positive specimen identification. Thermolyne says the double-walled insulation of the storage vessel reduces liquid nitrogen consumption, and the vessel's narrow neck lessens liquid nitrogen "boiloff". The \$1,525 (US) Locator 4 holds four racks of nine 81-ampule storage boxes, and the \$1,580 (US) Locator 8 stores eight racks of nine boxes holding 25 ampules each.

Growing ideas

LH Fermentation's diverse line of fermenters includes one that's easy on animal cells — the 502D slow direct-drive system (Reader Service No. 105). The 502D fermenter features a motor-driven agitator with a marine impeller that operates over the range 0–250 r.p.m. LH Fermentation says the fermenter's domed base-plate reduces shear forces and aids mixing at slow speeds. Vessels are available with 1.3 or 2.3 litre working volumes,

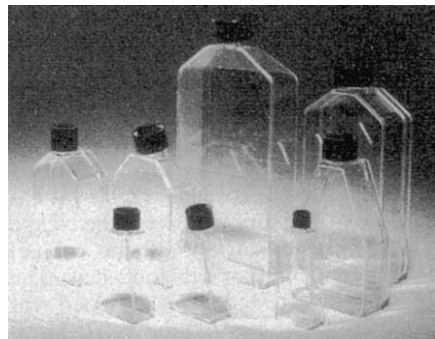


The marine impeller of the 502D fermenter makes it gentle enough for mammalian cells.

and are jacketed to prevent temperature fluctuations and eliminate the need for cartridge heaters that can cause localized overheating. The LH Fermentation ferm-

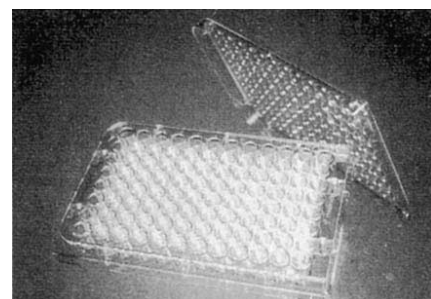
enters can be used for suspension cultures and cultures on microcarrier substrates. LH Fermentation will be in booth 704.

Costar, in booth 607, has recently introduced extra-large tissue culture flasks measuring 225 cm² (Reader Service No. 106). Costar's flasks have a 33 cm neck finish to maximize pipette access while



Large enough for any tissue culture need, Costar's new flasks have lots of room to grow in, minimizing contamination risk. They have a canted neck design, alpha numeric coordinates, and a high optical standard. The flasks are sold for \$86 (US) for 24.

Disposable, sterile 96-well tissue culture plates with lids are new from Dynatech Laboratories, in booths 4 and 5 (Reader Service No. 107). The plates are gamma-irradiated prior to packaging in a Secure-Seal pouch. If the seal on the package is broken, the pouch becomes cloudy to alert the user that the plate inside is no longer sterile. The plates feature flat-bottom wells of uniform thickness for optimum clarity, and alphanumeric codes for rapid well identification. Contamination between wells is reduced by the integral condensation rings on the one-way lid,



The Secure-Seal pouch packages for Dynatech's plates ensure sterility.

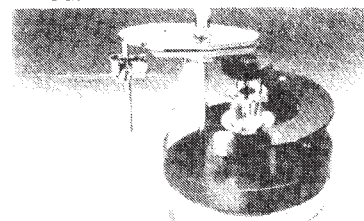
raised well rims and deep inter-well areas. The stackable plates cost \$93 for 50.

Magnifying the issue

Bausch & Lomb's StereoZoom 6 Plus microscope, in booth 12, was made with tissue culture needs in mind. It has new $\times 10$ wide-field eyepieces that permit a 21 mm field of view, to allow the operator to see more of the specimen and scan quickly at low power (Reader Service No. 108).

ADVERTISEMENTS

SINGER INSTRUMENTS MICROMANIPULATORS & ACCESSORIES YEAST? TETRAD ANALYSIS?

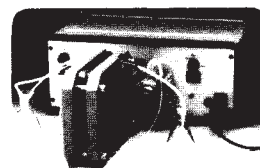


Singer MK III Micromanipulator and Precision Elevator.

Write direct to:
Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset TA23 0QL
England. Tel. (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50

LIPOSOMES REPRODUCIBLE — HOMOGENEOUS — EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.47

The lenses have special glass formulas and anti-reflection coatings for crisp, high-contrast images. The \$1,440 (US) microscope includes Bausch & Lomb's InstaZoom system, with integral adjustable stops, that provides the complete



StereoZoom 6 Plus with a 21 mm field of view.

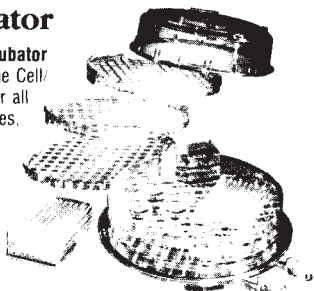
zoom range with less than one-third of a revolution of the control knob. The microscope has a 6:1 zoom ratio, a $\times 2$ - $\times 240$ magnification range and a parfocal zoom range of $\times 0.67$ - $\times 4.0$.

ADVERTISEMENTS

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

Low Cost



Eliminates evaporation loss and contamination. Easy to use - flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **billups-rothenberg, inc.,** pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.

Reader Service No.6

Medwire Div. stocks a wide variety of **TEFLON COATED***

Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.



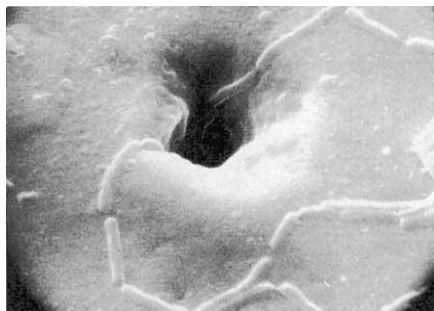
Medwire Div.

Affiliate of Sigmund Cohn Corp
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300

*Reg. T.M. of E.I. DuPont

Reader Service No.49

Advanced Biotechnologies, Inc., in booth 302, will be introducing its new **microanalysis/electron microscopy services** for the analysis of biological and materials specimens (Reader Service No. 109). The company's microscopy services



This Advanced Biotechnologies, Inc., micrograph shows bacteria on tooth enamel at $\times 6,200$.

include transmission EM, scanning EM, X-ray spectroscopy, freeze-etch techniques, thin-sectioning, negative staining and critical point drying. The new microscopy services complement ABI's existing services for mass cell culture, plasmid DNA purification and virus purification.

Protective hoods

The Baker Company, Inc. will be displaying its **SterilGARD laminar flow hood** in booth 6 (Reader Service No. 110). The SterilGARD hood is a Class II, Type A biological safety cabinet, making it suitable for use with low to moderate risk biological agents and minute quantities of toxic chemicals and radionuclides. SterilGARD has an 8 inch access opening with an average ambient air intake velocity of 100-110 f.p.m., and comes with one petcock and a warning alarm that sounds if the viewscreen is opened wider than the suggested 8 inches. All contaminated air inside the cabinet is surrounded by areas of negative pressure to prevent the escape of contaminated air should the cabinet become punctured or develop a leak. The HEPA filters, tested to be 99.9 per cent effective on particles greater than $0.3 \mu\text{m}$, according to The Baker Company, can be changed from the outside front of the cabinet. The Baker Company charges \$3,980 (US) for the 4-foot SterilGARD.

Labconco Corporation has just released a guide to their line of **Protector fibreglass hoods and accessories** (Reader Service No. 111). The 12-page catalogue includes a colour-coded chart to help researchers choose a fume removal system to fit their specific needs. The catalogue also compares the advantages of integral blower hoods vs. remote blower hoods, and spells out the specifications and dimensional data for Labconco's Protector-48, -60, and -72 hoods. Labconco's fibreglass hoods for international users are also described, highlighting the remotely-controlled cold water gooseneck faucet, the two service fixtures, and the choice of



LABCONCO

Rapid Selection Guide

for popular models of:
Protectors, Fibreglass Hoods and Accessories including Blowers,
Cabinets and Pumps

Labconco's guide helps whittle down the choices when purchasing a hood.

115-volt or 230-volt electrical configuration. The catalogue is available free from Labconco, exhibiting in booth 407.

Media market

A new **hybridoma propagation medium** is being offered by the American Type Culture Collection, exhibiting in booth 506 (Reader Service No. 112). Called Hybri-Care, the medium is chemically defined for the growth of a high percentage of hybridoma cell lines. Hybri-Care has a combined buffering system of HEPES and NaHCO_3 that enables it to be used in all stages of hybridoma production from fusion and subsequent cloning to single-cell subcloning procedures. Supplemented with 10 per cent fetal bovine serum, Hybri-Care can be used to improve the viability of difficult cell lines. Supplied in powder form, Hybri-Care costs \$10 (US) per pack to make 1 litre of reconstituted medium.

Flow has opened its vaults, and is selling its Gold and Silver Reserves **CELLect fetal bovine serum** products (Reader Service No. 113). CELLect Gold, suitable for hybridoma growth and monoclonal antibody production, is carefully tested to eliminate lots with even low levels of endotoxins. To earn the Gold label, lots of CELLect must have less than 1 ng ml^{-1} of endotoxins, no detectable mycoplasma contamination, and no BVD, PI, or IBR viruses. CELLect Silver is a standard-use serum, subjected to the same sterility tests as CELLect Gold, but not screened for endotoxins. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.